Group and individual evaluation in engineering project courses

Calin Stoicoiu
Conestoga College Institute of Technology and Advanced Learning
Cstoicoiu@conestogac.on.ca

Abstract - In engineering programs of study, students often work in small to medium size groups. In particular, programs designed on the project–based learning principle rely heavier on structured group work for projects and integrated courses. There are incontestable benefits surfacing from group work, particularly seen as the increase of critical thinking and problem solving skills and development of social interaction abilities.

Challenges occur more often than expected with the assessment and evaluation of the individual performance and participation of each group member in contrast with the whole group results. The final result can be outstanding but it might realistically belong to only one or two group members. The result of the individual assessment must not only reflect accurately and fairly one’s effort but also fit properly in the whole group diversity landscape. Alternatively, group members may have valuable but rather inconspicuous contributions that might easily be undetected and go unrecognized.

How to identify sooner rather than later the non-participating students within the group and correct the situation? Which is the best method of fair detection and praise for considerable contribution? The extraction of peer evaluation data and its incorporation in the overall group assessment, represents another often difficult or misinterpreted task. These are questions and challenges yet to be properly addressed. This paper provides a synopsis of existing evaluation techniques for engineering students working in groups, both from a psychological and academic point of view, including examples of current practices from existing project–based learning programs.

Keywords: Engineering, Individual, Group, Evaluation, Participation, Results, Peer, Review.

1. INTRODUCTION

As employers of engineering graduates increasingly demand the ability to work effectively in teams, collaborative learning and group based projects are now ever more integral parts of a modern engineering post-secondary education.

While the academic faculty is expected to assess the students individually, tying a student’s success to individual results, simultaneously, expectations to teach team skills and assess group work of their students are often at odds. Learning in groups has obvious potential benefits but at the same time produces the risk of group member masked contributions more often, non-contributions to the final result. Although team learning is effective to help students learn, the assessment of individual students in a team is not straightforward or effortless. To encourage teamwork, it is essential that students feel confident that they will be rewarded fairly for their contributions [1]. Weak contributors to the group’s results must be identified early in the project or course deployment and the proper actions must be taken for their self-correction.

There are numerous publications of studies and developed methodologies for group and individual evaluation, some relying on mathematical techniques to process the evaluation data, relying on factoring of evaluation criteria or the evaluating party (self, peer or faculty). The application of these diverse methods is hindered or hastened by a large number of criteria and particular environmental characteristics present throughout engineering programs:

• Number of students in the project course
• Group size
• Student-selected versus faculty-designated group membership
• Complexity of the project deliverables
• Faculty expertise and workload
• Project course timelines

This paper will look at few existing methods of evaluating individual and group performance, highlighting their advantages and shortcomings. The majority of methods rely on a certain form of mathematical modeling (analytical methods) but ultimately all of them are based on capturing an individual participation factor used to alter the team performance indicator.
The assessors are in all cases represented by one lead faculty or faculty team (responsible for assigning team and individual marks) and students themselves (providing self and peer evaluations).

It is observed that the analytical methods, although capturing the real experiences of over and under contribution, soft skills and theoretical knowledge, generally leave out any psychological criteria that, according to recent studies, are very important and influential for an accurate representation of the final result grade.

This paper will also present some of the main psychological criteria and findings responsible for adjusting the analytical results. Group members end up having functional roles within the team and, although some of these roles are demanded by the project, many are self-elected by the students within the group. Little direct attention seems to have been given to the training of group members in the membership and leadership roles required for effective group growth and production [3].

For comparison, we will look also at the group and individual evaluation system used for the third year project of the Mechanical Systems Engineering B. Eng. program at Conestoga College. Although this model follows some of the practices outlined in the presented methods, it has few unique characteristics that allow the students to contribute to project outcomes with their preferred skill set while being evaluated for their contributions considering certain psychological aspects of their activity, as observed by faculty throughout the course.

It is the intent of this publication to offer some alternatives to the existing group and individual evaluation techniques employed in engineering higher education.

2. OVERVIEW OF EXISTING EVALUATION METHODS

2.1 Individual contribution assessed by faculty

The initial focus is on evaluation methodologies employed by faculty and project supervisors. These methods rely on continuous contact between the group of students and faculty. Studies show that if evaluation relies on only the final course or project result (even if comprehensive) the accuracy of the individual mark will be at best average [6]. The ideal practice is to have weekly meetings with the group and to observe and record their weekly progress. The evaluation must rely on the entire semester performance covering a multitude of deliverables and milestones.

Nepal [1] produced a comparison of seven different analytical methods for capturing team and individual performance. The equations governing these methods are outlined in Figure 1. The described methods consider the individual contribution as a fraction from the team total, expressed as percentage, annotated as ICF (individual contribution factor), as follows:

$$ICF = \frac{\text{Individual contribution} (\%) }{100/n}$$

Where n = number of group members

It is understood that one of the first steps at the project start will be the group members’ selection and the delineation of expected tasks the group will be responsible for. Although these expected individual contribution shares will be known before the actual evaluation takes place, sometimes this task distribution is either difficult to establish clearly or will change throughout the project progression.

Some of the methods described in Figure 1 below create a more accurate reflection of the real performance than others, in particular methods 4 and 5. Summarizing the methods’ characteristics:

- Method 2.1 - doesn’t recognize additional contributions while penalizes under-achievement
- Method 3 – partially addresses individualistic behavior but is too generous with under contribution
- Method 3.1 - corrects the under contribution issue from Method 3
- Method 4 - generates lower individual marks than other methods using the highest achievement level as benchmark
- Method 5 - introduces a cap factor \( a \) to provide refined control over rewards and deductions however it is more difficult to implement.

<table>
<thead>
<tr>
<th>Method</th>
<th>Equation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Method 2.1</td>
<td>( f(ICF) = \begin{cases} \frac{ICF \text{ for } ICF \leq 1}{ICF} &amp; \text{for } ICF &gt; 1 \end{cases} )</td>
</tr>
<tr>
<td>Method 3 SPARK</td>
<td>( f(ICF) = \sqrt{ICF} )</td>
</tr>
<tr>
<td>Method 3.1 SPARK-Knee</td>
<td>( f(ICF) = \begin{cases} ICF \text{ for } ICF \leq 1 \ \frac{ICF}{ICF + 1} \text{ for } ICF &gt; 1 \end{cases} )</td>
</tr>
<tr>
<td>Method 4</td>
<td>( f(ICF) = \frac{ICF}{ICF_{max}} )</td>
</tr>
<tr>
<td>Method 5 Nepal</td>
<td>( f(ICF) = \begin{cases} ICF \text{ for } ICF \leq 1 \ ICF - \frac{ICF - 1)^2}{2a(1 - \frac{TM}{100})} \ 1 + (a/2) \left( 1 - \frac{TM}{100} \right) \end{cases} )</td>
</tr>
</tbody>
</table>

Figure 1

All methods will generate an individual mark \( IM \):

\[ IM = TM \times f(ICF) \]

Where:

- \( IM \) = individual mark
- \( TM \) = team mark
- \( ICF \) = individual contribution factor
- \( f(ICF) \) = multiplying factor, method specific
The first two methods presented by Nepal [1] are considerably weaker in capturing the individual performance and therefore are not shown here.

The summary above illustrates very well the advantages of introducing more control over the nature of real-life group-member contributions (under or overachievement), however with added complexity to the resulting calculations. One problem encountered in practice is that the higher the complexity of the method equations, the less likelihood of students to understand, accept and follow the evaluation scheme. If the evaluation method is not well understood, students will continue to doubt and often challenge its results.

The above examples calculate the individual marks derived from the group performance which is then modified by a multiplying factor. The various mathematical models define the correction factor $f$ (ICF) using an increasing degree of refinement. The main advantage of these methods is that they will generate a reasonably accurate final individual mark given by the professor, or project coordinator.

Their limitation, however, originates from the initial definition of the individual contribution as a fraction of the group result. For complex project courses having many specified deliverables, this approach will be inappropriate. Instead, separate group and individual marks should be generated and combined by an algorithm for the final course grade.

2.2 Individual contribution evaluated by group members

There are two aspects of group work that can be assessed: the product and the process [8]. We define the product as the concrete project results represented by analysis, calculations and documentation which are all based on the knowledge students acquire in the previous and current theoretical courses. In other words, the product is represented by project’s actual deliverables. Generally, professors are the group work product assessors and determine a grade based on using established sets of criteria for each project deliverable. This grade will reflect the group’s performance, attributed to all group members.

The assessment of the process is rather more difficult because much of the work occurs outside formal teaching sessions [8]. The process refers to all the activities, behaviors and actions the students in the group will undergo to arrive at the finished product(s). In many instances, after initial problem definition and data are provided by faculty, engineering students divide the workload and work on tasks by themselves. Often, only brief meetings will take place between the faculty or project coordinator and students. There are not many programs of study where students have significant dedicated project class hours weekly, supervised by the project faculty lead.

As a result, the process of evaluation will be passed on to the students themselves, as they hold the most relevant information of the project process and this becomes what is the current, well established method of peer evaluation. Peer evaluation is complemented in many instances by self-evaluation, although there are still conflicting opinions about the relevance of self-evaluation. Some researchers (Goldfinch 1994) have argued for the inclusion of self-assessment and others, for its exclusion (Falchikov 1993) [8]. There are extensive models, however, that use both, with no obvious differences between the results they generate. How do we incorporate peer and self-assessment into each student’s individual result? It seems that the most used model involves the generation of an Individual Weight Factor (IWF). This factor should be distinguished from the previously mentioned ICF factor. The ICF controls the individual mark from the product (performance) standpoint, originating from faculty, whereas IWF will indicate the process quantification (behavior and attitude), being in most cases student driven.

One example of a peer assessment model is described in where the IWF as defined above is generated using[7]:

$$IWF = \frac{IER}{AER}$$

Where:

$IER = $ Individual Effort Rating and represents the sum of all ratings received by a student form all group members for all given assessment criteria.

$AER = $ Average Effort Rating for the group calculated from all IER values.

The final mark awarded to an individual student is actually a mark-up or mark-down by factoring in an individual’s contribution reflected by the IWF [7].

All self and peer evaluation models, including this method, are affected by two factors:

1. The relevance of the assessment criteria given to group members to use in the evaluation.
2. Sincerity and realism of each student self-assessment as well as for their peers’.

The assessment criteria can be easily developed and optimized around the project scope and deliverables to collect relevant data, still, the reliability of the student assessors is much harder to control. Unreliable assessors (showing a large score deviation from the rest) can produce false results. The iterative IWF method (it-IWF) uses the student evaluator reliability as a modifying factor and through a series of iterations, calculates the individual contribution score up to a satisfactory results convergence.

Because the main concept of the it-IWF is to calculate the reliability of each evaluator, an unfair and biased evaluator can be identified and their impact diminished [8]. This method appears complex and less user friendly however it can be easily implemented as a macro running in readily available spreadsheet software packages.
created and tested, it will produce fast and effective results.

It is not the intent of this paper to describe the method in detail, however, an example of the file containing this macro source code as well as all the examples in this paper is available upon request to the author [8].

Possible limitations of this method are:
1. The student assessors having low reliability can express their true and honest perceptions; this occurrence could be caused by conflicting roles or where the student assessor and the evaluated student have incompatible personalities.
2. As a relatively new method, more extensive use of the it-IWF method is required to test its effectiveness.

3. PSYCHOLOGICAL ASPECTS

3.1 Group-Individual Interdependence

It is largely recognized that having students organized in groups for accomplishing larger scale tasks has incontestable advantages. Real life activities starting at family level and ending with professional activities require us to interact and cooperate with other people, as all of us have limitations in what we can achieve alone. Even though people are capable of living separate and apart from others, they join with others because groups meet their psychological and social needs [5]. As we start interacting with others, we will experience affinity, diversity and sometimes rejection. Due to our uniqueness in thinking patterns, visions and aptitudes, we need to learn how to recognize ideas dissimilar to our own and promote courageously our firm beliefs even though both scenarios can be subject to positive or negative reception by others.

The main benefit of group associations comes from the accumulated result of all the unique contributions from individuals and for the group to survive and thrive through the process, it must reach equilibrium. This operating equilibrium is attained when certain functional roles within the group are assumed and recognized. Another criterion is the recognition of the effect that the group itself will have over each. Research has identified a few common requirements that contribute to recognition of a group: interdependence, social interaction, perception as a group, commonality of purpose, and favoritism [4]. The group influence on individual behavior can manifest in positive and negative ways, and this is clearly visible in team work settings within engineering projects as well. Positive effects are detected easily; a favourable group dynamic will generate good results, as each individual will contribute constructively and almost effortlessly to the final product or outcome. Studies show however that there are negative group influences at work as well. These negative developments are more difficult to discover and usually their effects are noticed later in the activity, when corrections are difficult to implement. There are three key negative influence processes worth mentioning: group think, group shift and de-individualization.

3.1.1. Groupthink: psychological phenomenon manifested within a group of people that transforms the desire for conformity and homogeneity of group results into shifted and deviant individual decision making patterns. Individual motivation and drive for original and unique ideas becomes engulfed and minimized by the group unanimity and emphasis on cohesion. Looking at group projects in engineering education, the evaluation of individual performance becomes burdened and blurred by the strong desire of an appropriate team result, sacrificing some unheard and potentially great ideas belonging to less vocal students. This will cause loss of individual creativity and undermine independent thinking. Add to this situation a strong group leader acting aggressively, dominating and seeking recognition and the dynamic of that group can deteriorate albeit they may still show a superior project result in the end.

3.1.2. Groupshift: phenomenon in which the initial positions of individual members of a group are exaggerated toward a more extreme position [4]. One group member in isolation will often have a conservative approach to risk and its magnitude as we tend to act more carefully and assess risk more cautiously when alone. There isn’t anyone around us to lean on, seek advice from, or to encourage us. A group scenario brings with it the sense of additional support and help and as a result, an increased perception of safety that naturally leads to accepting more risk. We abandon the fear, reduce the doubt and become more active when in a cohesive group, however if pushed over normal limits, these initial individual boosts transform into excess and exaggeration.

The theory behind this shift is that the group dynamic allows the individual members to feel that their position is correct or supported, and they will feel more comfortable taking on more extreme views, as other members of the group support their initial ideas [4]. With groupshift, a false impression of support for more extreme ideas and decisions builds in each group member with the end results being exaggerated and extreme actions that disregard the real risks.

In engineering projects, this effect will be recognized through apparently revolutionary and very non-conformist group generated ideas that seem extremely original and enticing but often bear dangerous implementation risks. If a general approval for these outcomes is detected, it is time to intervene and have the group perform a risk assessment combined with the proposal of alternate, safer solutions.

CEEA16; Paper 28
Dalhousie University; June 19-22, 2016 - 4 of 8 -
3.1.3 De-individuation: the drastic reduction of individuality. People experiencing this trait will become lost in the group, agreeing to everything the group decides and pursues. Certain group members will be at risk of suffering from de-individuation. Non-participation (passivity), total approval for group decisions (conformity) and systemic favorable responses to group’s requests (compliance) are all components of de-individuation but must be distinguished from it when observed alone.

The loss of individuality has very significant effects on results of engineering group projects as it manifests by a sharp decrease in creativity and originality of ideas. The healthy debate, questioning other’s opinions and the demand of reasoning for any negative comment, all diminish with a visible drop in quality of results.

If these effects are not detected early and firmly corrected, the individual evaluation results will be skewed with the loss of optimal outcomes for the group. Students that show a general lack of participation, always approving of faculty or other team member’s ideas, are just drifting away taken by the group current. They will most likely have a low individual mark when assessed by faculty and peers.

3.2 Group Members’ Roles

A group draws its healthy dynamic and smooth functioning from two sources: the group leader providing cohesiveness and guidance and the properly designated group members’ roles. Usually a lot of emphasis is placed on the leadership exercised by the group leader represented by the professor or program team assuming this role. Groups may operate with various degrees of distribution of “leadership” functions among group members or of concentration of such functions in one member or a few individuals [3]. Meanwhile, the other specific roles and their effects are generally neglected. “Membership” is tacitly identified with “followership” [3].

Roles assumed by members within a group are often confused with their individual personalities. One can play a specific role influenced by one’s personality thus showing a certain amount of inter-dependence between the two; however, member roles influence the group behavior, whereas members’ personalities remain at the individual level. Groups in different stages of an act of problem selection and solution will have different role requirements [3]. Similarly, students in a project group setting will take on different roles appropriate to the task at hand. For example, one student can be an “information seeker” during the research phase and an excellent “energizer” later on when the group faces an impasse.

A classification of group member’s roles is presented in [3] as they belong to three categories:

1. Group task roles – contribute to the selection, definition and solution of the project’s tasks.
   Examples: initiator-contributor, information seeker, opinion giver, energizer, recorder.

2. Group building and maintenance roles – responsible for assuring the group’s normal functioning
   Examples: encourager, harmonizer, gate-keeper, standard-setter.

3. Individual roles – not group oriented, serve strictly individual goals which are not relevant to the group’s interests. Must be detected early and sanctioned if could compromise group’s results.
   Examples: aggressor, blocker, recognition-seeker, dominator.

We must recognize roles that belong to the three categories as they will define the type of individual contribution and will influence in a positive or negative way the group’s final result. In practical terms, the faculty project leader must observe how these roles are assumed by the students in the group and how they morph and adapt as the project challenges will require it. Neglecting them places the group dynamic at risk and affects the detection of the individual students’ contributions.

Peer and self evaluation tools provide the main avenue to control the group’s process. By using relevant assessment criteria these will identify the group members’ roles and their effects.

4. EVALUATION PRACTICES IN MSE

In the Mechanical Systems Engineering program at Conestoga College, group and individual evaluations are utilized mainly within the yearly project courses. The Canadian Engineering Accreditation Board (CEAB) prescribes a list of hard as well as soft skills which a graduate should acquire throughout a program [11]. In this program the third and fourth year project courses are designed for the students to develop several CEAB graduate attributes and therefore require more sensitive methods to capture and process collective and individual work. The third year project consists of two separate courses, Part A and Part B, offered in adjacent academic terms and each utilizing different methods of evaluation.

4.1 Marking methods

Part A course of the third year project has structured deliverables which are assessed by the project coordinator using detailed marking rubrics with task-specific criteria and weightings toward the final mark. Students are assessed for performance (product) as well as behavior and attitude (process). If teamwork skills are listed outcomes of the course, one needs to ensure that the performance, behavior and attitude of the individual
members of the team are enhanced by undertaking the course activities [11].

The evaluation of the individual participation relies on three sources:
1. Project coordinator’s (faculty) notes for each group member’s participation
2. Students’ records in their project logbook
3. Peer and self-evaluation data

This process will produce three marks reflecting the group result as well as individual contributions, as shown in Figure 2.

Each of the M1, M2 and M3 marks has an associated weight which varies from task to task. Some tasks require a solid group result (product) while others rely more on individual performance or behavior/attitude (process). Higher M1 weight will emphasize the process while M2 can capture psychological aspects of the process. M3 will define the attitude and behavior at the student level. The final individual student mark M for a particular task, can be calculated as in this example:

\[ M = M_1 (w_1) + M_2 (w_2) + M_3 (w_3) \]

Where:
- \( w_1, w_2, w_3 \) are fractional weights determined based on importance of the marks M1, M2, M3 and must sum 1.

Appendix A1 illustrates examples of assessment tools used to generate the final mark after incorporation of both group and individual evaluation attributed to each student.

Contrasting to activities in Part A in the first half of the year, those in Part B are mainly focused on practical skills and deliverables. Not being as closely tied to theory, the implementation of the project is not as clearly defined anymore. The evaluation becomes more ethereal, more difficult to define or control in minute detail. Students are applying their preferred skill sets and are evaluated on their specific contributions. The goal is that the entire class, divided in functional groups completes the project build. Each student receives marks proportional with the level of completion of the chosen tasks. An example of this approach is illustrated in Appendix A1, Figure A3.

It must be specified that project course tasks and milestones have different types of deliverables such as: written reports, CAD models and drawings, computer programs, electrical and pneumatics diagrams, just to name a few. It is therefore important to have appropriate evaluation tools for each deliverable. Multiple rubrics are utilized having specific criteria fitting each task.

As the project course has four hours of contact weekly with the students in each of the two academic terms, the faculty coordinator acts as a mentor and tutor, offering support and guidance. The course follows the flipped classroom model, where students complete the work under faculty supervision. In addition, all student teams have meetings with the coordinator, presenting their progress and discussing problems or challenges they encounter. During these meetings faculty can easily observe details around students’ group roles, or any effects of groupthink, groupshift or lack of individual presence among the students, acting as mediator when such manifestations occur or offering suggestions on self-mediation. This demonstrates how psychological factors are integrated in the evaluation of the process component (behavior, attitude) of the project course.

4.2 Observed limitations and proposed improvements

Although the MSE project courses evaluation methodology has proved capable of capturing reliably the individual product and process details, including some psychological group characteristics, a few aspects in need of improvement have been identified:

- Faculty driven individual assessment must contain correction factors for under- and over-achievement.
- Peer and self evaluation data must be refined by correcting unfair and biased assessments.
- Identification and categorization of project deliverables and tasks that are based on the prevalence of performance, behavior and attitude components.
- Development of peer evaluation and faculty assessment rubrics need to include criteria for measuring the group psychology effects.

For Part B project course, the evaluation of the individual performance can be complemented by introducing a faculty-led process assessment component, as currently this component is completed through peer evaluation alone. Faculty would have to record any observations based on rubrics that contain criteria meant to capture the individual student’s behavior and attitude characteristics. A proposed rubric for faculty-led individual psychological process assessment is shown in Appendix A1, Figure A5. This rubric will be introduced in the coming academic year for both year 3 project courses. The rubric uses has two major components: individual and group induced psychological factors [3]. Sub-components A.1 and A.2 cover the positive individual traits constructive to the group while A.3 and component B capture the negative influences. Marks between 0 and 5 are allocated, all negative factors having negative values. The total score is calculated with a weight of 80% for individual factors and 20% for the negative group effects.

The proposed rubric for group process assessment will be tested for the upcoming academic year and will be amended based on observed effectiveness.
4. CONCLUDING REMARKS

There is no doubt that there are numerous group and individual evaluation techniques implemented across engineering programs of study. The complexity of data processing algorithms varies with the level of the assessment precision and fairness sought. Methods relying on complex mathematical processing are more difficult to understand and accept however deliver more accurate results.

Project courses raise additional challenges as they have multiple, different and often changing deliverables. It is important to recognize that both product (performance) and process (behavior, attitude and individual in a group psychology) must be captured in the evaluation of individuals. Professors are experts at evaluating the product but the process component assessment needs additional effort. Some of the avenues to implement it are: increased contact time between instructor and students, utilizing task-specific rubrics and criteria, records of meetings and a periodic progress reporting system.

Effects of group and group members’ psychology should not be neglected as these are often responsible for the group dynamic and individual progress. There are no established solutions to capture the psychological facets of the group environment, however studying the existing published material around this topic and recognizing the identified problems definitely helps with the development of concrete detection and processing tools or rubrics.

CEAB graduate attributes covering the soft skills are more difficult to implement and measure, however project courses offer the perfect setting to achieve this. Teamwork, Communication skills, Project Management, Professionalism are attributes that engineering students master through projects. The only thing remaining to do is to correctly and completely evaluate students’ activity when working in groups and acting as individuals, detecting behavioral and group psychology negative effects while attentively observing each student’s involvement in, and reaction to, the group structure.

References


APPENDIX A: EXAMPLES OF EVALUATION TOOLS IN MSE

A.1 Project deliverable group evaluation criteria

A.2 Calculation of the deliverable final mark

A.3 Peer evaluation questionnaire

A.4 Task-specific individual performance evaluation

A.5 Proposed process evaluation rubric