Evaluating Individual Contributions in Teams for Project-Based Learning in Data Science

Arko Barman
Data to Knowledge Lab
Rice University
Houston TX, USA
arko.barman@rice.edu

Su Chen
Department of Statistics and Data Science
The University of Texas at Austin
Austin TX, USA
su.chen@austin.utexas.edu

Genevera I. Allen
Department of Statistics
Columbia University
New York NY, USA
genevera.allen@columbia.edu

***Abstract*—In this Innovative-Practice Full Paper, we present the design and implementation of a novel strategy for the assessment of the individual contributions of students working in teams for an interdisciplinary project-based learning (PBL) course. Team-based and project-based learning have been the cornerstones of experiential learning in recent years. In such courses, the assessment of students is often performed solely at a team level, resulting in student discontent, concerns about fairness, and social loafing. The application of a single method for evaluating individual contributions has its shortcomings, prompting us to incorporate a combination of evidence-based approaches. Thus, to mitigate these problems, we developed an assessment strategy that involves three components – self and peer evaluation (SPA), instructor's evaluation of individual contribution, and class participation. Anonymized SPA was carried out at three different checkpoints over the semester for both formative and summative assessment. The individual contribution of the students is also evaluated by the instructor and/or other mentors based on their interactions with all the students in the team over the entire semester. Class participation is another component of individual contribution incorporated in our scheme, where instructors evaluate the participation of individual students in classroom activities, presentations, and meetings. We collected feedback about the perception of students towards our assessment policies for individual contribution. We observed overall satisfaction and positive attitudes toward our scheme. We noted positive student perceptions of fairness in grading through our scheme and reduced chances of social loafing. Further, in the feedback, the students noted the effectiveness of using SPA as an evaluation tool, the usefulness of instructor's evaluation, and the role of class participation in creating a more engaging classroom and a more enriching experience. We also emphasize the need to set clear expectations for students with regard to the grading policies early in the semester. Finally, we discuss our experiences in implementing our design and recommendations for incorporating our grading scheme in team-based PBL courses.**

***Index Terms*—student assessment, collaborative learning, project based learning, experiential learning, active learning**

I. Introduction

Project-based learning (PBL) is a form of active learning developed using the constructivist theory in education, where learners actively participate in the process of education through experiential learning. Especially when combined with team-based learning, it fosters learning through a social process through the context that every project creates [1]. Working in teams also promotes teamwork skills and the students acting as self-regulated learners [2]. Capstone courses offer the ideal backdrop for PBL in teams. Such courses have been developed for business [3]; science, technology, engineering, and mathematics (STEM) [4], [5], and in particular, computing and data science education curricula in several universities with further enhanced experiential learning components such as real-world client-facing projects [6]–[8].

While such a course encourages social interdependence through cooperation as student team members share common objectives and the outcomes of individual students are affected by their teammates [9], it has been widely acknowledged that there is a need for promoting individual accountability and implementing a mechanism for fair assessment of individual students [10], [11]. In fact, students often have negative perceptions if the same grade is assigned to all members of the team [12]. Having a grading policy that adequately weighs in individual contributions also results in reduced "social loafing" (individuals putting in lower effort compared to their teammates) and "freeriding" (an individual reaping the benefits of their team's output while contributing less than others) [13]–[16].

Peer assessment and feedback is possibly the most widely adopted approach used by instructors as a proxy for the evaluation of individual contributions to a team project [17]–[21]. Besides creating a framework for fair assessment of individuals, it also encourages personal reflection, creates opportunities for developing professional skills, and greater satisfaction with the course and grading [17], [21]. Student perceptions of the peer assessment process are positive in general, especially when the peer assessment is anonymous [22], [23]. However, there are drawbacks to the peer assessment strategy. Entrusting students with a complex task like the assessment of individual performance in a group might be unreasonable, especially since it is a particularly challenging task even for experienced instructors [24]. Students also often lack the confidence to judge peers, might feel socio-emotional discomfort, and develop a sense of distrust [17], [21], [25]. On the other hand, instructors perceive that students are inexperienced to make such judgments [25].

An alternative strategy is the incorporation of a self-assessment component tied with the peer assessment [26]–[30]. Self and peer assessment (SPA) emphasizes a reflection

on one's individual contributions, thus adding a perspective to peer assessment that is absent in peer assessment by itself. However, this is not the perfect solution as students often tend to rate their own contributions higher than their teammates [28]. There were also observed cases of "benign collusion", where all team members agreed to give themselves and each other the same ratings, both in the literature as well as in our experience [31], thus rendering the entire process useless. Overall, it has been acknowledged that there is a lack of empirical studies on peer assessment and self-assessment [11], [22].

Thus, a combination of different assessments is perhaps the best choice while evaluating the contributions of individual students in a team project [27]. Besides SPA, we also considered studies on other assessment methods such as instructor's evaluation of individual contribution. In fact, studies have shown that students often favor the instructor's evaluation over SPA for assessing individuals, placing more trust in the instructor's experience and ability to make judgments with regard to grading [24], [32]. However, we noted a gap in the literature on whether the student perception of individual contribution matches the assessment by the instructor and whether the students feel that the instructor spent enough time with the team to understand and gauge individual contributions.

Further, we also considered the aspect of including the assessment of individual students based on their participation in classroom activities, such as presentations and team meetings with the instructor and/or mentor. Although assessment based on classroom participation is well-studied in lecture-style classrooms [33]–[35], we noted a lack of literature on this aspect of grading from the perspective of team- and project-based learning. We also noted a lack of studies discussing student perceptions of how class participation affects grading.

We describe our grading strategy for a PBL course in data science, which is a combination of team and individual assessments for the final grade of each student. Although developed for a data science course, we envisage that the grading strategy can be used across all disciplines in any PBL course. In developing our framework for individual assessments, we incorporated three components – (i) self and peer assessment (SPA); (ii) instructor's evaluation of individual contributions; and (iii) participation of individual students in the classroom activities, e.g., presentations and team meetings involving the instructor and/or mentor. The rationale behind this implementation is to minimize the risks associated with each method of assessment while reaping the benefits of all of them. To evaluate the effectiveness and student perception of our grading scheme, we collected feedback from the students through a questionnaire that included responses on a Likert scale as well as free-text responses.

The contributions of our study are summarized below:

- For the assessment of individual contributions to a team project, we implemented a novel assessment strategy that includes three components – self and peer assessment, instructor's evaluation of individual contribution, and class participation.
- We collected and analyzed feedback on student perceptions of our grading scheme.

II. Course Structure & Grading Overview

In this section, we briefly provide background on the PBL data science capstone course that is the focus of this study as well as outline our grading and assessment scheme and its implementation.

A. Data Science Capstone Course

In 2018, we designed and launched a novel interdisciplinary and client-sponsored capstone program in data science. We briefly give an overview here to explain the context of our grading mechanisms and this study. In this PBL course, undergraduate, master's, and occasionally Ph.D. students from a variety of disciplines tackle a semester-long client-sponsored data science or machine learning project. This course serves as the required capstone for undergraduate and master's degrees in data science as well as an option to fulfill capstone requirements for eight other degree programs. Potential data science projects from client sponsors in industry, research, and government or non-profits are framed and set up before the semester begins and students entering the course can rank their preferences for available projects. Based on student preferences as well as their background, skill sets, and other demographic factors, students are assigned to interdisciplinary teams. Given that students do not have the option to choose their own teammates, team and individual assessment strategies are particularly critical for the success of this course.

Throughout the semester, students work on their course projects with the guidance of a Ph.D. student mentor as well as the course instructors. Teams are required to meet weekly amongst themselves and also with client sponsors. Class time is used in a hybrid manner. At the beginning of the semester, instructors lead lectures and discussions on several critical aspects of data science: the data science pipeline, ethics, teamwork, and communication. For the remaining two-thirds of the semester, class time is used in part for interim presentations and check-ins from each team as well as team-working sessions and additional time where instructors can mentor each team. At the end of the semester, each team gives a final presentation of their work and submits a final report and software product in the form of a Git repository. This data science capstone program has been running at our university for six years with 667 students successfully completing a data science project to date.

B. Grading & Assessment

We developed a grading and assessment scheme that accurately captures the quality of the team-based projects as well as fairly balancing individual efforts by incorporating a variety of individual and team assessments.

a) Grading Overview: To balance team-based and individual-based assessments, we set the overall grading breakdown as follows:

- Team: 70%

  - Interim Evaluations: 15%
  - Final Report: 25%
  - Final Presentation: 15%
  - Final Software: 15%
- Individual: 30%
  - Self & Peer Assessment (SPA): 10%
  - Instructor & Mentor Assessment: 10%
  - Class Participation: 10%

Note that with 30% of the grade based on individual assessments, students who do not contribute to the project or are "social loafing" [14] can earn up to three letter grades lower than their teammates (e.g. teammates earn an "A" and an individual can earn a "D").

b) Team-Based Assessment: Team-based assessments are based on the quality of the team's final report, final presentation, and final software product. At the beginning of the semester, students are provided with detailed grading rubrics and grading checklists for each of these three deliverables. The grading rubric for the final report, for example, includes a breakdown of 30% for communication, 35% for technical data science approach, 15% for data science reproducibility, 10% for significance and impact, and 10% for creativity and innovation. Similarly, the presentation rubric consists of five equal parts: organization and structure, oral communication, visual communication, general audience communication, and technical communication. Finally, our software rubric has a breakdown with five parts: 25% for readability and organization, 25% for documentation, 20% for computational reproducibility, 20% for data science reproducibility, and 10% for efficiency. The rubrics will be made available by the authors on request.

In addition to these detailed grading rubrics, we provide several examples of excellent ("A" or "A+" range) reports, videos of teams' final presentations, and Git repositories with instructor comments to set expectations for the course and provide positive examples of successful projects as well as illustrate how the grading rubrics are implemented. Permissions from each student and the client sponsor were obtained. We also provide a select few "B" and "C" range examples of final products with instructor comments that are illustrative of some common pitfalls that teams encounter.

c) Individual Assessment: Self & Peer Assessment (SPA): While many self and peer evaluation schemes have been proposed [27], [29], we use the CATME system [36] with questions on a Likert scale measuring as well as free-responses.

Importantly, SPAs are due three times throughout the semester: an initial, midterm, and final evaluation. The initial evaluation is due about a month into the semester, the midterm evaluation is done halfway through the semester, and the final evaluation is due at the end of the semester. The mean of these three assessments is used to compute the scores used towards 10% of the final grade. Instructors also pay careful attention to improvements that each student makes based on the initial SPA. The SPA requires every individual to rate themselves as well as their peers for the following metrics on a 5-point Likert scale – contributing to work, interacting with teammates, keeping the team on track, expecting quality, and having knowledge/skills. In addition, every team member is encouraged to provide free text responses evaluating themselves and their teammates, with the express purpose of providing constructive criticism so that their teammates can improve. The responses are anonymized and released for every team after the response period has ended.

Additionally, we require every student to respond to the following questions that are not used for grading (also on a 5-point Likert scale) – team satisfaction, team cohesiveness, and psychological safety. The feedback from these questions is evaluated by the instructors to identify potential conflicts within the team and other problems in teamwork that might arise within each team. These responses are also released after anonymization.

Besides the responses that are released to the teams, the students are also provided an option to provide free text comments to the instructors that are confidential and not released to the team. In our experience, students utilize this option mostly to express concerns regarding non-contributing teammates and conflicts within the team. Thus, this also acts as a form of confidential communication between every individual student and the instructors and is viewed as an effective tool for the instructor to understand team dynamics and resolve conflicts within a team as soon as they are reported.

Note that the initial and midterm SPA are useful as formative assessments so that students can take heed of the peer feedback and improve their contribution to teamwork, commitment, and engagement if needed. The final SPA is summative in nature and useful in evaluating the contribution of all members to the teamwork as perceived by their peers vis-à-vis their self-assessment.

d) Individual Assessment: Instructor Evaluation: Instructors and team mentors evaluate each student's individual contribution to the team project. These evaluations come from observations of the students during the class, mentoring sessions, individual contributions acknowledged in the team's report, and individual contributions to the software product documented in the team's Git repository. It is to be noted that training sessions and lectures are provided in the course to familiarize students with version control and Git. In weekly mentoring sessions in the classroom, students present on their progress each week and the team is supposed to hold each student accountable. Hence, mentors can carefully assess each student's contribution as well as their progress toward accomplishing shared team goals. Importantly, we also require that all students be involved in coding and data analysis aspects of the project and hence we carefully check the Git commit history to understand which technical software portions each student has contributed.

e) Individual Assessment: Class Participation: Both instructors and team mentors evaluate class participation for each student. This includes (i) recording observations on attendance and active participation in class, team meetings, and team-sponsor meetings; and (ii) evaluating individual contributions

to in-class and final presentations. During discussions on teamwork at the beginning of the semester, students are instructed to actively participate in class and team meetings; students are also taught best facilitation practices to help each team member speak up and contribute in team meetings. Students are also required to present at least one interim or update presentation in class as well as participate in the final team presentation. Thus, instructors and mentors have several opportunities to assess each student's individual participation in team presentations.

III. Methods & Study Design

Throughout the course, responses to survey questionnaires are collected from the students. For example, at the beginning of the semester, student demographics information was collected in the course application survey, along with a questionnaire evaluating their background in different areas of data science, such as statistical analysis, computer vision, natural language processing, time series analysis, and signal processing, as well as their preferences for the available sponsored projects in that semester. The student responses for this survey are used by the instructors to group students into teams and assign every team to the most suitable project based on the background and preferences of the students. Similarly, an exit survey is conducted to understand student perceptions of the course and possible directions for improvement.

Another survey conducted at the end of the semester is designed to evaluate our implementation of the grading scheme for assessing individual contributions through feedback collected from the students. The survey questionnaire used for collecting feedback evaluates the student perceptions of the overall individual assessment strategy as well as for the three components separately. This survey mostly consists of questions requiring answers on a Likert scale but it also has the option for the students to submit free text responses. Our study, data collection, and analysis were approved by the university institutional review board (IRB).

All questions on a Likert scale have 5-point responses, where 1 corresponds to "strongly disagree" and 5 corresponds to "strongly agree".

A. Feedback on Overall Assessment Strategy

Feedback for the overall assessment strategy for individual contributions consists of the following questions and a section for comments.

- A.1 The proportion of individual (30%) vs. team (70%) grading is appropriate.
- A.2 Grading on individual contribution results in more fairness in grading for every team member.
- A.3 Grading on individual contribution forces everyone to contribute to the best of their abilities and prevents team members from slacking.
- A.4 I am satisfied overall with individual contribution grading policies.

B. Feedback on Self and Peer Assessment

Student perceptions and feedback on self and peer assessment grading policies were collected using the following questions along with a section for comments.

- B.1 Questions for peer evaluation are appropriate for gauging individual contribution.
- B.2 Anonymization of peer evaluation helped me express the contributions of my teammates more freely.
- B.3 Proportion of grading based on peer evaluation (10% of final grade) is appropriate.
- B.4 Peer evaluation provided constructive criticism for improving my skill set, understanding and overcoming my shortcomings, and improving my contribution to the team.

C. Feedback on Instructor's Assessment of Individual Contribution

The policies on the instructor's assessment of individual contribution were evaluated using the following questions and a section for comments.

- C.1 The instructor/mentor has sufficient interaction with the team to understand and assess every member's individual contributions.
- C.2 I think that the instructor/mentor's assessment of individual contributions is similar to your experience of working with your teammates.
- C.3 Proportion of grading based on the instructor/mentor's assessment of individual contribution (10% of final grade) is appropriate.
- C.4 Discussions with the instructor/mentor helped me improve my individual contribution to the team.

D. Feedback on Assessment based on Class Participation

Assessment of students based on class participation was evaluated using the following questions in the survey along with a section for comments.

- D.1 Grading on class participation ensured that all team members participated in the activities and the mentoring sessions during the class.
- D.2 Greater class participation from all students results in a more engaging classroom and a more enriching experience.
- D.3 Proportion of grading based on class participation (10% of final grade) is appropriate.
- D.4 Grading based on class participation improved my participation in the classroom.

E. Student Demographics

To provide more context for the study design, we summarize the student demographics information below. Tables I - IV show the breakdown of students enrolled (by both count and percentage) by academic level, gender, race/ethnicity, and major/minor/academic department. For Table IV, only the first major listed by a student was considered, while a student might have had multiple majors and minors.

TABLE I
DISTRIBUTION OF ENROLLED STUDENTS BY ACADEMIC LEVEL: UNDERGRADUATE, MASTER'S, AND PH.D.

Level	Count	Percentage
Undergraduate	85	51.2%
Master's	72	43.4%
Ph.D.	9	5.4%

TABLE II
DISTRIBUTION OF ENROLLED STUDENTS BY GENDER. NOTE THAT OTHERS COMBINES "NON-BINARY" AND "PREFER NOT TO SAY" CATEGORIES.

Gender	Count	Percentage
Female	59	35.5%
Male	104	62.7%
Others	3	1.8%

IV. RESULTS

In this section, we present student perception of the grading scheme, which is the main result of this study. The survey collecting feedback on grading schemes for assessing individual contributions has been running for three semesters starting Fall 2022. The enrollment number over these semesters is 166, and a total number of 138 survey responses have been recorded, with a response rate of 83%.

A. Quantitative Results

The aggregated results of all survey questions except the free-text responses are shown in Fig. 1 along with percentages of responses in each of the 5 values on the Likert scale and mean numerical score for each of the questions. Most questions on overall assessment and SPA received overwhelmingly positive responses, indicating that the students were satisfied with the overall assessment strategies as well as SPA policies in our course design.

The question that received the most positive responses was Question D.2: Greater class participation from all students results in a more engaging classroom and a more enriching experience. This response clearly shows the importance of promoting students' participation and engagement in the classroom. The implementation of grading schemes that can assess a student's engagement in a team-based learning environment is crucial. However, the challenge is how to assess students' participation in a transparent and fair manner, which could be the main focus of improvement.

On the other hand, the question that received the least positive response was Question C.4: Discussions with the instructor/mentor helped me improve my individual contribution to the team, which indicates the instructors could have spent more time on assessing team dynamics, formulating strategies for facilitating more effective discussion and providing clearer guidelines for individual accountability within teams. Overall, students responded with positive feedback on the grading scheme.

TABLE III
DISTRIBUTION OF ENROLLED STUDENTS BY RACE/ETHNICITY. NOTE THAT OTHERS COMBINES "MIXED" AND "PREFER NOT TO SAY" CATEGORIES.

Race / Ethnicity	Count	Percentage
Asian	102	61.4%
Black / African American	4	2.4%
Hispanic / Latino	8	4.8%
White	41	24.7%
Others	11	6.6%

TABLE IV
DISTRIBUTION OF ENROLLED STUDENTS BY DEPARTMENT/MAJOR/MINOR.

Major	Count	Percentage
Computer Science	71	42.8%
Statistics	39	23.5%
Electrical & Computer Engineering	25	15.1%
Data Science	13	7.8%
Other STEM	12	7.2%
Social Sciences	6	3.6%

B. Qualitative Results

In addition to the quantitative results on the survey questions, a small number of students provided free-text comments to the following questions. Some meaningful comments are summarized below as themes with the corresponding question.

- Do you have any comments on the current policies or suggestions to improve the overall policies?
  - Suggestions were made for more weight on the overall individual contribution (especially in teams with a high variance of perceived contribution, i.e., one team member significantly under- or over-contributes).
- Do you have any comments on the current peer evaluation policies or suggestions to improve the policies?
  - CATME self and peer assessment process was lengthy and time-consuming.
  - Concerns were raised by students regarding anonymity and constructive feedback from the peer evaluation for smaller teams.
- Do you have any comments on the current policies of assessing individual contributions by the instructor/mentor or suggestions to improve the policies?
  - Some students recommended more time from the instructor/mentor on team mentoring to assess individual contribution.
- Do you have any comments on the current policies of assessing class participation or suggestions to improve the policies?
  - Concerns were raised by students regarding transparency and fairness of assessment: introvert students being put at a disadvantage and the need for better methods to promote meaningful engagement.

Based on these comments, we will discuss plans for improving the effectiveness and fairness of grading individual contributions in this course to better support student learning and team dynamics.

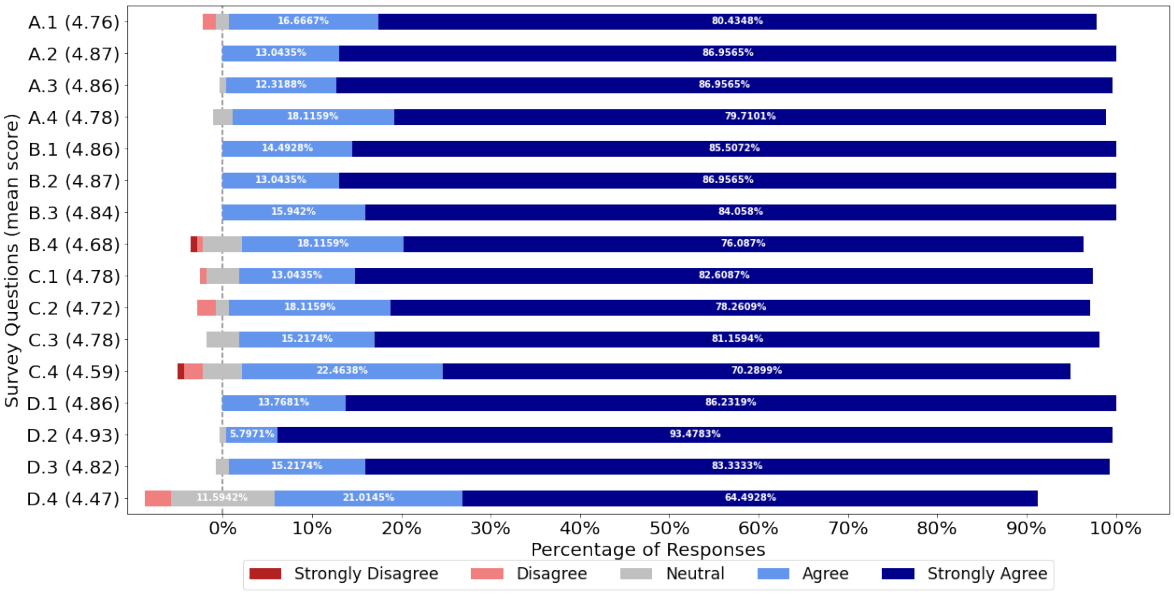

Fig. 1. Visualization of the survey responses on a Likert scale of 5 responses - "Strongly Disagree", "Disagree", "Neutral", "Agree", and "Strongly Agree"..

V. Discussion

Several strategies for evaluating the contribution of individual students to a team have been reported in prior studies, with peer evaluation being the most common. Although it has been acknowledged that a combination of different strategies is the best way to evaluate individual contribution in teams [27], we noted a lack of studies that report a combination of assessment strategies.

In our experience of designing and implementing the combination of three evidence-based techniques for assessing individual contributions to teams, we have found our strategy to be very successful. In particular, the success of our strategy is supported by the perception of students that we report in our collected feedback, especially responses to A.4 (overall satisfaction with the individual contribution grading policies) where 97.8% of respondees expressed their satisfaction while the remaining were neutral. We also note that important components of our design, such as the proportion of individual vs. team grading (97% positive response to A.1) were very well received by the students. Further, it is important to note that our study reaffirms the findings in prior studies that emphasize greater fairness in grading when incorporating an individual assessment component in the grading (100% positive response to A.2).

In our implementation of SPA, we observed that students expressed entirely positive perceptions regarding the appropriateness of the SPA questionnaire (responses to B.1), anonymity in peer feedback (responses to B.2), and the proportion of grading allocated to SPA (responses to B.3). This is a testament to the effectiveness of the CATME system [36] and validation of prior studies on the use of anonymity in SPA [22], [23]. However, students expressed slightly mixed opinions (albeit not negative) about whether they received constructive criticism in the SPA process. In the future, we wish to make students aware of the role of constructive feedback in formative assessment using SPA and some examples of how to provide constructive criticism.

Student perceptions of the instructor's evaluation of individual contribution (responses to C.1 - C.4) were slightly more mixed compared to the other components. Based on the feedback we received, we would like to explore more ways for the instructors and mentors to interact with the team, have a better understanding of the student perceptions of their teammates' contributions, and most importantly, have more one-on-one discussions with individual students on how to improve their contributions to the teamwork.

From our analysis of the survey results, we noted that class participation is the component of individual assessment that the students expressed the most displeasure about. While all students felt that this component ensured that all team members participated in the activities and mentoring sessions (responses to D.1) and this led to a more engaging and enriching classroom experience (responses to D.2), their own participation did not necessarily improve due to this grading component (responses to D.4). This observation is somewhat similar to prior studies where it was reported that students tend to rate their own contributions more than those of others in SPA [28].

In particular, from the free-text responses, we observed that some students might be shy and less likely to participate actively in presentations and meetings. Some students also expressed discontent about feeling pressured to present and

participate in the class due to this grading component. Keeping this in mind, we wish to explore, evaluate, and incorporate alternate methods and strategies to encourage the participation of students in team meetings, mentoring sessions, and presentations.

Overall, with respect to the assessment of individual contribution, providing clear explanations of these grading policies early in the semester is crucial. For example, students should be aware of how SPA affects their grades and also its formative assessment role for initial and midterm evaluations. Similarly, for the instructor's assessment, the instructors should clearly communicate to the students how their participation in the in-class mentoring sessions, commits to the team's Git repository, and participation in the progress reports play a role in evaluations. With regard to class participation, students should be made aware that their attendance, participation in team meetings and sponsor meetings, and involvement in the interim and final presentations affect their grades. In our course, these expectations and grading policies were also reinforced and codified through the use of a team contract that every team must draft and adhere to throughout the semester.

Similarly, for the team grading components, we strongly emphasize setting clear expectations for the students early in the semester with respect to the grading rubrics. To set the standards of the deliverables, it is also important that the students are provided examples of reports, presentations, and software that received good grades in previous semesters.

Anecdotally, we also noted some innovative grading strategies that preclude the involvement of the instructor and mentor in the grading of the team deliverables. Some universities implement a policy that requires deliverables to be graded by external evaluators (other faculty members in the department who are not the instructors) to remove any biases in grading that might occur if the instructor grades the teams.

We want to note here that as the surveys were completely anonymized, there was no possibility to analyze the responses vis-à-vis demographic information and team membership for particular teams. This approach was designed to ensure that the students responded without any inhibition. To aid in more fine-grained and detailed analysis, in the future, we aim to incorporate a voluntary option for respondents to provide demographic information and team membership details if they wish.

VI. Conclusion

In this study, we report the design and implementation of a novel grading strategy for the assessment of individual contributions of students to their teams in an interdisciplinary client-facing PBL course. Besides grading team deliverables, we used a combination of self and peer assessment, instructor's evaluation, and class participation to award grades for individual contributions of students. We collected feedback on student perceptions of our grading strategy and observed generally positive responses from students, thus indicating that our scheme was well-received by students and successful in the course. In particular, students expressed overwhelmingly positive views with respect to fairness in grading and overall grading policies. We also share our experiences and recommendations about implementing this grading scheme.

Acknowledgment

GA acknowledges support from NSF DMS-2210837.

References

[1] L. Helle, P. Tynjälä, and E. Olkinuora, "Project-Based Learning in Post-Secondary Education – Theory, Practice and Rubber Sling Shots," *Higher Education*, vol. 51, no. 2, pp. 287–314, Mar. 2006.

[2] B. J. Zimmerman, "A social cognitive view of self-regulated academic learning," *Journal of Educational Psychology*, vol. 81, no. 3, pp. 329–339, 1989.

[3] A. Gilinsky and R. Robison, "A Proposed Design for the Business Capstone Course with Emphasis on Improving Students' Information Competency," *Journal of Management Education*, vol. 32, no. 4, pp. 400–419, Aug. 2008, publisher: SAGE Publications Inc.

[4] S. Howe and J. Goldberg, "Engineering Capstone Design Education: Current Practices, Emerging Trends, and Successful Strategies," in *Design Education Today: Technical Contexts, Programs and Best Practices*. Cham: Springer International Publishing, 2019, pp. 115–148.

[5] A. J. Dutson, R. H. Todd, S. P. Magleby, and C. D. Sorensen, "A Review of Literature on Teaching Engineering Design Through Project-Oriented Capstone Courses," *Journal of Engineering Education*, vol. 86, no. 1, pp. 17–28, Jan. 1997, publisher: John Wiley & Sons, Ltd.

[6] R. F. Dugan Jr, "A survey of computer science capstone course literature," *Computer Science Education*, vol. 21, no. 3, pp. 201–267, Sep. 2011.

[7] P. Anderson, J. Bowring, R. McCauley, G. Pothering, and C. Starr, "An undergraduate degree in data science: curriculum and a decade of implementation experience," in *Proceedings of the 45th ACM technical symposium on Computer science education*, ser. SIGCSE '14. New York, NY, USA: Association for Computing Machinery, Mar. 2014, pp. 145–150.

[8] G. I. Allen, "Experiential Learning in Data Science: Developing an Interdisciplinary, Client-Sponsored Capstone Program," in *Proceedings of the 52nd ACM Technical Symposium on Computer Science Education*, ser. SIGCSE '21. New York, NY, USA: Association for Computing Machinery, Mar. 2021, pp. 516–522.

[9] D. W. Johnson and R. T. Johnson, "An Educational Psychology Success Story: Social Interdependence Theory and Cooperative Learning," *Educational Researcher*, vol. 38, no. 5, pp. 365–379, 2009.

[10] D. W. Johnson and R. T. Johnson, *Assessing Students in Groups: Promoting Group Responsibility and Individual Accountability*. Corwin Press, Dec. 2003.

[11] J. Forsell, K. F. Frykedal, and E. H. Chiriac, "Group Work Assessment: Assessing Social Skills at Group Level," *Small Group Research*, vol. 51, no. 1, pp. 87–124, Feb. 2020, publisher: SAGE Publications Inc.

[12] R. L. Barfield, "Students' Perceptions of and Satisfaction with Group Grades and the Group Experience in the College Classroom," *Assessment & Evaluation in Higher Education*, vol. 28, no. 4, pp. 355–370, Aug. 2003.

[13] A. R. Jassawalla, A. Malshe, and H. Sashittal, "Student Perceptions of Social Loafing in Undergraduate Business Classroom Teams," *Decision Sciences Journal of Innovative Education*, vol. 6, no. 2, pp. 403–426, 2008.

[14] I. Fronza and X. Wang, "Towards an Approach to Prevent Social Loafing in Software Development Teams," in *2017 ACM/IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM)*, Nov. 2017, pp. 241–246.

[15] C. Gabelica, S. De Maeyer, and M. C. Schippers, "Taking a free ride: How team learning affects social loafing," *Journal of Educational Psychology*, vol. 114, no. 4, pp. 716–733, 2022.

[16] U. Samarakoon, K. Manathunga, and A. Imbulpitiya, "Enhancing the Quality and Student Attitude on Capstone Projects Through Activity-Based Learning and Reduced Social Loafing," in *Computer Supported Education*, ser. Communications in Computer and Information Science. Springer International Publishing, 2022, pp. 336–358.

[17] A. Burgess, C. Roberts, A. S. Lane, I. Haq, T. Clark, E. Kalman, N. Pappalardo, and J. Bleasel, "Peer review in team-based learning: influencing feedback literacy," *BMC Medical Education*, vol. 21, no. 1, p. 426, Aug. 2021.

[18] M. Lejk and M. Wyvill, "Peer Assessment of Contributions to a Group Project: A comparison of holistic and category-based approaches," *Assessment & Evaluation in Higher Education*, vol. 26, no. 1, pp. 61–72, Jan. 2001.

[19] P. L. Kench, N. Field, M. Agudera, and M. Gill, "Peer assessment of individual contributions to a group project: Student perceptions," *Radiography*, vol. 15, no. 2, pp. 158–165, May 2009.

[20] D. R. Lane, "Peer Feedback Processes and Individual Accountability in Team-Based Learning," in *Team-Based Learning in the Social Sciences and Humanities*. Routledge, 2012, num Pages: 12.

[21] C. M. Cestone, R. E. Levine, and D. R. Lane, "Peer assessment and evaluation in team-based learning," *New Directions for Teaching and Learning*, vol. 2008, no. 116, pp. 69–78, 2008.

[22] E. Panadero and M. Alqassab, "An empirical review of anonymity effects in peer assessment, peer feedback, peer review, peer evaluation and peer grading," *Assessment & Evaluation in Higher Education*, vol. 44, no. 8, pp. 1253–1278, Nov. 2019.

[23] B. Sridharan, M. B. Muttakin, and D. G. Mihret, "Students' perceptions of peer assessment effectiveness: an explorative study," *Accounting Education*, vol. 27, no. 3, pp. 259–285, May 2018.

[24] K. F. Frykedal and E. H. Chiriac, "Assessment of students' learning when working in groups," *Educational Research*, vol. 53, no. 3, pp. 331–345, Sep. 2011.

[25] C. Stefanou, S. Stefanou, and L. Hood, "Feedback and Change: Assessment of Individual Contribution in Collaborative Activities in the Higher Education Classroom," *Journal of Excellence in College Teaching*, vol. 12, pp. 77–91, Jan. 2002.

[26] L. E. Willcoxson, ""It's not Fair!": Assessing the Dynamics and Resourcing of Teamwork - Lesley E. Willcoxson, 2006," *Journal of Management Education*, vol. 30, no. 6, pp. 798 – 808, 2006.

[27] F. Dochy, M. Segers, and D. Sluijsmans, "The use of self-, peer and co-assessment in higher education: A review," *Studies in Higher Education*, vol. 24, no. 3, pp. 331–350, Jan. 1999.

[28] L. Johnston and L. Miles, "Assessing contributions to group assignments," *Assessment & Evaluation in Higher Education*, vol. 29, no. 6, pp. 751–768, Dec. 2004.

[29] Z. Ma, X. Yan, and Q. Wang, "Assessing individual contribution in collaborative learning through self- and peer-assessment in the context of China," *Innovations in Education and Teaching International*, vol. 57, no. 3, pp. 352–363, May 2020.

[30] P. J. Pastirik, "Using problem-based learning in a large classroom," *Nurse Education in Practice*, vol. 6, no. 5, pp. 261–267, Sep. 2006.

[31] M. R. Fellenz, "Toward Fairness in Assessing Student Groupwork: A Protocol for Peer Evaluation of Individual Contributions," *Journal of Management Education*, vol. 30, no. 4, pp. 570–591, Aug. 2006.

[32] J. Alden, "Assessment of Individual Student Performance in Online Team Projects," *Journal of Asynchronous Learning Networks*, vol. 15, no. 3, pp. 5–20, Jun. 2011.

[33] C. Holly, S. Porter, T. R. Vitale, and M. Echevarria, "Grading participation in the classroom: The assumptions, challenges, and alternatives," *Teaching and Learning in Nursing*, vol. 19, no. 1, pp. 27–33, Jan. 2024.

[34] L. A. Paff, "Does Grading Encourage Participation? Evidence & Implications," *College Teaching*, vol. 63, no. 4, pp. 135–145, Oct. 2015.

[35] K. E. Czekanski and Z. R. Wolf, "Encouraging and Evaluating Class Participation," *Journal of University Teaching and Learning Practice*, vol. 10, no. 1, 2013.

[36] A. C. Loignon, D. J. Woehr, J. S. Thomas, M. L. Loughry, M. W. Ohland, and D. M. Ferguson, "Facilitating Peer Evaluation in Team Contexts: The Impact of Frame-of-Reference Rater Training," *Academy of Management Learning & Education*, vol. 16, no. 4, pp. 562–578, Dec. 2017, publisher: Academy of Management.