

Digital Peer Grading in Group Learning: Empirical Insights and Best Practices

Eduard Klein
Business Department
Bern University of Applied Sciences (BUAS)
Bern, Switzerland
e-mail: eduard.klein@bfh.ch

Claus Noppene
Business Department
Bern University of Applied Sciences (BUAS)
Bern, Switzerland
e-mail: claus.noppene@bfh.ch

Abstract—The impact of digital peer grading on group learning dynamics in project-based courses in higher education is studied using the web-based tool Peer Grading Tool (PGT). Following an action research approach, this study, which was conducted over two years across 20 courses, examines the use of PGT for both grade influencing and reflective peer assessment. Empirical data collection from students and lecturers led to interesting findings. The results reveal nuanced insights into the contextual suitability of PGT, highlighting factors, such as group size, familiarity among members, heterogeneity, voluntariness, and robust feedback culture. The research also highlights potential challenges and emphasizes the importance of a pre-existing feedback culture to mitigate negative impacts. The findings contribute to the derivation of actionable recommendations and best practices for the implementation of peer grading campaigns in higher education.

Keywords—peer grading; group learning; peer assessments; free riding; group work reflection.

I. INTRODUCTION

In practice-oriented higher education, group work is often employed to facilitate independent and problem-oriented learning, as it effectively prepares students for project-oriented and agile work environments [1]. Ideally, all team members contribute equally to the overall success of the project, albeit often in individually diverse ways [2][3].

Peer assessment and peer grading are both methods of evaluating a peer's work, but they differ in their purpose and process. Peer assessment is a formative approach that aims to help students plan their learning, identify strengths and weaknesses, and develop personal and professional skills [4]. It involves students providing feedback to their peers about their work. On the other hand, peer grading is a more evaluative process where students assign grades or scores to their peers' work [5].

Peer grading is also a widely adopted method for group members to reflect on collaboration and evaluate individual contributions to the overall outcome. Through feedback from peers, individual learning opportunities are unlocked. Peer grading also serves to counteract undesirable free riding, which contradicts fairness principles and undermines the motivational functions of performance assessments.

Resulting from peer grading are dual outcomes; firstly, it serves as a mechanism for reflective practices within group learning processes. Secondly, the outcomes serve as a basis for individualized grading decisions. This can be through a *reflective grading* approach (see Section II), wherein peer feedback impacts the group process without directly impacting individual grades. Alternatively, an *individual grading* approach, aligned with the *assessment of learning* concept [6], involves using peer grading results to deviate grades from the group result potentially.

This duality underscores the nuanced role of peer grading in shaping not only individual grades but also the broader learning experience within collaborative group contexts. By adopting this didactic approach, lecturers harness the potential of peer grading to promote collaborative learning, encourage self-reflection, and drive iterative improvements in both group processes and outcomes.

Peer grading also unlocks valuable learning opportunities while simultaneously avoiding unwanted social loafing. The mere expectation of negative feedback from peers triggers positive behavioral changes in learners [7]. Free riding poses a central challenge in group-oriented learning contexts [8], as performance-oriented students perceive group work with free riding as an overall frustrating experience [9].

The rest of this paper is organized as follows. Section II provides an overview of our research design. In Section III, the peer grading tool PGT is described. Section IV presents the results of accompanying research on how the use of digital peer grading in group learning and presents best practices. The conclusions in Section V close the article.

II. RESEARCH DESIGN

Following an action research approach, we study the use of digital peer grading and its implications using the web-based *Peer Grading Tool* (PGT) [10]. Our goal is to contribute to the improvement of teaching in higher education: How can digital peer grading enhance learning in groups for students? How can it improve our teaching and coaching of group work as lecturers?

Empirically, we report on data generated over 4 consecutive semesters (Q4/2021-Q2/2023) in 20 different courses using the peer grading tool. The data collection was conducted using both qualitative and quantitative methods

(survey, focus groups) and involved students as well as lecturers.

Our research covers different types of assessment as well as of peer grading. Depending on the didactic setting, there are several types of assessments. *Assessment of learning* focuses on measuring students' knowledge and skills, often through tests and exams [11]. In contrast, *assessment as learning* involves students in the assessment process, promoting self-regulation and decision-making [12]. This approach aligns with the concept of *assessment for learning*, which emphasizes using assessment to improve learning and provide feedback [13][14].

Two types of peer grading, *reflective* and *individual*, were used in this study. The *individual* approach is employed at the end of group work and falls into the category of assessment of learning. The *reflective* approach is also applied during group activities and falls into the category of assessment as learning or for learning.

In this study, the individual approach was also used to individualize group marks. To clarify, the term "marks" refers to the deliverables assessed by the lecturer, while the term "grade" refers to the point scale of the peer assessment criteria. For each group member, the deviation from the group average of peer grading is calculated (see Section IV), and thresholds for these variations are established. If a group member's deviation from peer grading surpasses an upper (or lower) threshold, defined for a course, an improvement (or deterioration) of the individual mark compared to the group mark can be made. This method allows for the individualization of group marks.

The surveys used structured as well as unstructured questions and covered a broad range of relevant topics (e.g. prior experiences of groups learning; attitude towards group learning, group dynamics or work attitude, etc.). In particular, the survey focused on the respondent's experience of peer grading. The following list of statements serves as a sample item. The respondents were asked to indicate their approval of the following statements on a 5-point scale (1 = "completely disagree"; 5 = "completely agree"):

- (1) Peer grading has made free riding more difficult.
- (2) Work was faked in the group to receive a good grade in peer grading.
- (3) Peer grading made it possible to reflect on one's own role or group behavior.
- (4) Peer grading increased the pressure within the group.
- (5) In peer grading, students tried to evaluate actual performance.
- (6) Peer grading makes the module assessment fairer.
- (7) In addition to actual performance, students also considered irrelevant criteria (e.g. sympathy) when grading their peers.
- (8) Peer grading enables students to work better in upcoming group work.
- (9) Peer grading makes it easier to recognize the performance of all group members.
- (10) Peer grading strengthens the feeling of trust in the team.

Section IV discusses the outcomes of the evaluations in more detail.

III. THE PEER GRADING TOOL PGT

The search for suitable peer grading tools tailored to our teaching contexts led to the evaluation of several systems, which we examined based on our specific requirements. It is noticeable that comparable peer grading tools are predominantly available through commercial licensing models, exemplified by Purdue University's CATME [15], or have limitations such as dependency on participants having a Google account, as with a peer grading tool integrated into Google Spreadsheets [16], which contradicts intended principles of flexibility, openness, and free accessibility.

The Moodle activity *Workshop* has similarities with the planned peer grading scenario. However, the analysis revealed that key features, encompassing functionality, granularity, flexibility, and customization, are not available in the Moodle activity *Workshop*. Therefore, the decision was made to implement the PGT tool as an in-house development and use Angular technology to create a publicly accessible web-based application.

We considered the following key requirements:

- (1) Each group member can grade each other group member (peer assessment) and also herself (self-assessment) according to defined criteria based on a numeric scale.
- (2) The criteria (names) and the numeric grade scale are definable by the lecturer during the creation of a peer grading campaign. The number of criteria (currently up to 10) and the criterias' weights are definable.
- (3) The criteria are not necessarily focused on some artifact or group product. Instead, they could also focus on the group process or project management aspects.
- (4) In case some or all criteria focus on an artifact, there is no upload required (as forced by some other tools). Instead, artifacts are communicated or shared outside the tool, e.g., via a learning platform.
- (5) The criteria catalog contains a description of each criterion's grade to enable objective and equal grading for all groups. By decision, it cannot be imported into the tool to avoid the complex handling of the possible variety of semi-structured documents. Instead, it must be provided outside the system.
- (6) Participants' data, such as name, e-mail address, etc., are provided as CSV data for bulk import. The tool provides an appropriate template.
- (7) Privacy and data protection are enforced: after a peer grading campaign, each student receives system-generated feedback containing the deviation of his self-assessment grade from the peer assessment grade. The peer assessment grade for student x is calculated as the average of his peers' grades concerning student x .
- (8) To facilitate tool use, students don't have to register, but instead receive a system mail with a personalized link to his or her grading form (see Figure 1).
- (9) Lecturers must register since they carry out a workflow and must come back into their tool session from time to time.

Peer grading form for Everdeen Katniss

campaign name: Project Management Lecture 2024
 group number: 1
 group size: 4
 grading scale: 1 - 8
 campaign start: Jan 31, 2024

criterion (weight)	Teamworking skills (1)	Quality (1)	Quantity (1)	Reliability (1)	average
Klein Eduard	8	3	7	4	5.5
Dent Arthur	4	3	4	8	4.75
Everdeen Katniss	3	8	6	7	6
Parsons James Joseph	3	8	8	4	5.75

comments
 comment for Arthur Dent
 well done!

Submit

Figure 1. Peer grading form for Katniss Everdeen.

The creation of a peer grading campaign by the lecturer is shown in Figure 2, demonstrating flexibility in creating and editing criteria.

edit campaign

in general

campaign name: Project Management Lecture 2024
 maximum number of points per criterion: 8
 language: english
 status: created
 created on: Jan 26, 2024

grading criteria

name of criterion:
 weighting:

name	weighting	delete
Teamworking skills	1	delete
Quality	1	delete
Quantity	1	delete
Reliability	1	delete

overview of group

download template | import list of participants | new group

number	number of participants	delete
1	4	delete
2	3	delete

group 1 | delete group
 group 2 | delete group
 save

Figure 2. Creating a peer grading campaign.

Peer grading evaluation for Everdeen Katniss from Apr 2, 2024

campaign name: Project Management Lecture 2024
 group number: 1
 group size: 4
 grading scale: 1 - 8
 campaign period: Jan 31, 2024 - Feb 3, 2024

criteria (weighting):
 Teamworking skills (1)
 Quality (1)
 Quantity (1)
 Reliability (1)

	Teamworking skills (1)	Quality (1)	Quantity (1)	Reliability (1)	average
peer assessment	3.67	3.67	5	4.33	4.17
self assessment	3	8	6	7	6

4.17

4.81

-0.64

peer assessment | group average | deviation

comments received

author	comment
Arthur Dent	Working with you was great.
James Joseph Parsons	It was a pleasure.

export overview

Figure 3. Peer grading result for Katniss Everdeen.

Upon completion of a peer grading campaign, all participants and the lecturer receive evaluations via system-generated e-mails. Each student receives anonymized feedback (see Figure 3) concerning the grading from his peers, and optional textual feedback as comments. The lecturer receives each other's gradings and the detailed gradings from and to each other's students (see Figure 4). Figure 3 shows the peer grading results for Katniss Everdeen, presenting her peers' grading per criterion, her self-grading, and (the big numbers from left to right) the average of her peers' gradings, the group average (=average of all peer gradings), and her derivation from the group average.

overview of group 1

grading from \ grading of	Klein Eduard	Dent Arthur	Everdeen Katniss	Parsons James Joseph
Klein Eduard	4.25	2	3	6.5
Dent Arthur	4.25	2.75	5.5	5.25
Everdeen Katniss	5.5	4.75	6	5.75
Parsons James Joseph	6	5.25	4	4.25
peer grading average:	5.25	4	4.17	5.83

comments from Arthur Dent:
 for Eduard Klein: I was quite isolated in this group
 for Katniss Everdeen: Working with you was great.

comments from Katniss Everdeen:
 for Eduard Klein: such a cool group!
 for Arthur Dent: well done!

comments from James Joseph Parsons:
 for Eduard Klein: the team building pub crawl event was super cool!
 for Katniss Everdeen: It was a pleasure.

Figure 4. Lecturer's view of how group members graded each other.

The lecturer has control over the peer grading process at any time (see Figure 5). The functionality comprises editing campaign details, reminding defaulting students, closing campaigns, and generating detailed result views.

The software is available as open-source under MIT license at <https://github.com/digital-sustainability-lab/peer-grading-tool-mirror>.

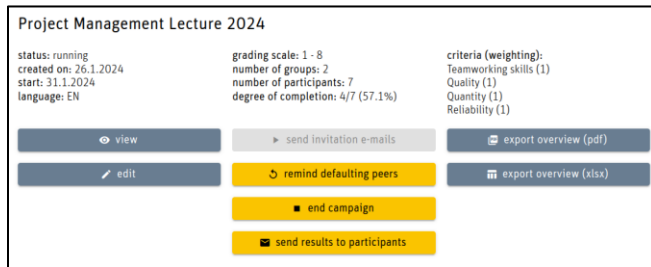


Figure 5. The lecturer has full control over the peer grading process.

Concerning the criteria catalog, for each criterion, the meaning of a grade must be defined. Guidelines for developing appropriate criteria are outside the focus of this paper. Instead, we list the grades' meanings for the criteria *ability to work in a team* for one of our study modules which has been proven appropriate for many years:

8/7: Exceptional contribution to the group process; "doer"; proactively advances project goals; is extremely committed; "sacrifices" himself/herself for project success; accomplishes much more than expected; opinion leader; high communication skills.

6/5: Committed member, contribution not too high, not too low; participates in the group process with an average sense of responsibility; has an integrating effect in conflicts.

4/3: Provides contributions on demand; neutral attitude towards project success; does not make motivational advances when there are "sags" in the team.

2/1: Unproductive to counterproductive attitude; does not contribute to the success of the project or hinders/prevents its progress. Destructive charisma.

In our study modules using peer grading, the criteria catalog is provided on the learning platform and discussed before the start of a peer grading campaign.

IV. RESULTS AND EXPERIENCES FROM PEER GRADING

The section covers three levels of results. Initially, the experience with individual grading is considered where the result of peer grading might influence the individual marks of the students. Second, results from the accompanying research are discussed, and finally, best practices are discussed.

A. Individual Grading

Figure 6 shows the results of a typical peer grading in a course. The x-axis represents the deviation of each student's grade compared to the group grade (which is the average of all peer grades in the group).

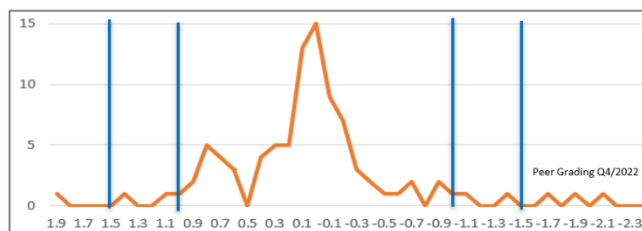


Figure 6. Results from one peer grading campaign. The x-axis shows deviations from the group grade, the y-axis shows the respective number of students.

The y-axis shows the number of students per deviation step. The blue lines define (arbitrary) thresholds at which the deviation leads to an individualized mark. The distribution shows that less than 10% of the students have a high or even extreme deviation. An analysis of the grades reveals that no discrimination based on gender or other diversity parameters can be found. Also, when comparing different campaigns, it is noticeable that the distribution is not influenced by the type of peer grading (i.e., reflective, or individual grading).

Occasionally, group-internal agreements such as identical peer gradings can be observed in the data. Although this may seem like a subversive practice that undermines the didactic concept, it is perfectly acceptable, as the individualization of marks should be limited to rather extreme cases. However, there is the scenario that multiple free riders in one group can potentially coordinate their grading activities and even downgrade potential outperformers.

Regardless of these group dynamic effects, our conclusion and a strict recommendation are: In the case of individual grading, the results of peer grading should not automatically lead to a change in the individual mark (see Figure 6). Instead, high deviations in peer grading should be compared with the lecturer's impression of the group and trigger a specific debriefing session with the group. This aligns with the fact that students sometimes tend to rate their peers' performance differently than their lecturer does and that peer evaluations may not always accurately reflect performance [17].

B. Experience with Peer Grading

The accompanying research on the experiences of students and lecturers (including surveys and focus groups) provides important insights.

In the students' feedback, several critical observations emerged in the free text fields. Firstly, there were concerns about the perceived fairness of the assessments. Some students expressed dissatisfaction with assessments that they felt were unfair and highlighted the need for transparency and consistency in grading.

In addition, students pointed out the influence of social factors such as likability or familiarity among fellow students, which they felt could have a subjective impact on the evaluation of fellow students. Finally, students shared their concern that peer grading could draw too much attention to the assessment tasks and thus undermine the learning experience.

However, the few consistently critical opinions expressed in the answers to the open questions are not confirmed by the quantitative results. The following statements receive high and very high approval rates.

- In peer grading, students tried to evaluate actual performance.
- Peer grading has made free riding more difficult.
- Peer grading makes the module assessment fairer.

The following statements continue to receive moderate approval:

- Peer grading made it possible to reflect on one's own role or group behavior.
- Peer grading increased the pressure within the group.

- In addition to actual performance, students also considered irrelevant criteria (e.g., sympathy) when grading their peers.
- Work was faked in the group to receive a good grade in peer grading.

Several of these statements have been derived from research findings, with our focus centered on the students' emphasis placed on statements related to peer grading. As per [17], peer grading can indeed make free riding more difficult, as it can lead to more accurate and reliable assessments. Furthermore, peer grading has been shown to facilitate reflection on one's role in group work [18].

Accordingly, important goals of peer grading can be achieved: Making free riding more difficult facilitates and promotes group learning. In addition, the final assessment is perceived as fair. Since our approach is not an experiment but action research, it is difficult to capture or even measure the effects of peer grading. Therefore, we asked students to compare their experiences with group work in different courses.

The analysis of the responses from the follow-up surveys among students yielded interesting findings:

- On average, peer-graded group work received comparable or better grades than other group work in all courses observed.
- Students acknowledge that peer grading reduces free riding, but this effect also leads to divergent opinions. It is crucial to recognize that, similar to an echo chamber, individual participants reinforce negative feelings and can thus have an unfavorable influence on the formation of opinions in the group.
- A small proportion of around 10-25% are against the continued use of peer grading.
- Conversely, around 75-90% are in favor of the continued use of peer grading. The distribution of the surveyed frequencies (rarely, occasionally, often, always) varies considerably depending on the course.
- A small proportion of 7-13% of students are in favor of the use of peer grading in all group work.

A correlation between rather negative statements regarding the use of peer grading and weak peer grading results of individual persons could not be investigated, as the surveys were conducted anonymously. For the same reason, it is not possible to determine which group of people took part in the surveys about peer grading and in what proportions (response rate approx. 30% across all courses). In the debriefing discussions, however, it became obvious that opinion leaders often come from a group of people who either contributed little to the group work and/or had poor feedback in the peer grading.

C. Best Practices

There are several contextual factors to consider when introducing digital peer grading. Peer grading seems particularly appropriate when student groups tend to be larger. This practical insight is in line with the fact that once a group is formed, group size influences the decisions of members to contribute to the group's public good. In a small group

members can easily notice if a member does not equally contribute to the group's efforts. However, as group size increases, free riding becomes more probable [19]. In this case, digital peer grading can compensate for the decreasing noticeability and perceptibility.

Another contextual factor to be considered is the familiarity of the team members. In scenarios where group members have limited knowledge of each other, peer grading can provide a fairer and more impartial assessment framework.

In addition, the heterogeneous composition of groups can benefit from peer grading as it allows different perspectives to be considered in the assessment. In cases where group formation is less voluntary, such as assigned or structured groups, peer grading provides a method of impartial assessment.

In educational settings where intense collaboration and self-organization are expected, peer grading also fosters a sense of ownership and responsibility among students.

Finally, a developed feedback culture helps when digital peer grading is used. In this case, digital peer grading offers new learning opportunities and serves as an acceptable method to reflect on the collaboration. If the feedback culture is not part of the didactic setting the use of digital peer grading might provoke dysfunctional group behavior and undermine group learning.

In summary, digital peer grading turns out to be suitable if the following characteristics are present:

- larger groups (>3 members)
- less familiar
- more heterogeneous
- less voluntary composition
- intense collaboration, self-organized
- higher feedback culture

Both the survey results and the analysis of the evaluations prove to be quite stable. They hardly deviate from the experiences documented in this section, not even for peer gradings in new classes.

V. CONCLUSION AND FUTURE WORK

Awarding identical marks in group work to all members of a group is only considered fair if all group members contribute to the group result in roughly the same way. In this study, peer grading is used on the one hand as a way of nuancing the group mark and on the other hand to reflect on the group work. To minimize the workload for lecturers and students concerning peer grading, a digital peer grading tool is used that automatically compares self-assessment and peer assessment and calculates individual deviations of group members from the group average.

The experiences and best practices from Section IV have so far been repeatedly observed in new peer grading courses, which is a sign of stability for best practices. A larger study is planned for the future, which will carry out extended evaluations and analyses about a possible segmentation of courses into groups of courses with similar characteristics (study program, group work skills etc.). This will be possible because new courses from other disciplines will be added as

the PGT tool becomes more widespread. Specific experiences and best practices may emerge, differentiated by segment.

Extensions are also planned concerning the (multilingual) peer grading tool. PGT currently runs as a single server instance including a database in which all user data is stored. For data protection and performance reasons, use with third-party hosted containers and the possibility for own branding will also be offered, including analytics for an overview of the use of installed instances.

ACKNOWLEDGMENT

The PGT tool was developed with funding from the BUAS Virtual Academy (Bern University of Applied Sciences). The study was conducted in the context of the DigiPeG Project [20] which has funding from the BeLEARN association which is an initiative of the Canton of Berne.

REFERENCES

- [1] A. S. Burke, "Group Work: How to Use Groups Effectively," *The Journal of Effective Teaching*, vol. 11, pp. 87–95, 2011, [Online]. Available: <https://api.semanticscholar.org/CorpusID:265889597>
- [2] D. B. Kaufman, R. M. Felder, and H. Fuller, "Peer ratings in cooperative learning teams," *ASEE Annual Conference Proceedings*, no. June, pp. 3931–3942, 1999, doi: 10.1.1.193.23.
- [3] S. Brutus and M. Donia, "Improving the effectiveness of students in groups with a centralized peer evaluation system," *Academy of Management Learning and Education*, vol. 9, no. 4, pp. 652–662, 2010, doi: 10.5465/AMLE.2010.56659882.
- [4] K. J. Topping, "Peer Assessment," *Theory Pract*, vol. 48, no. 1, pp. 20–27, 2009.
- [5] H. Li *et al.*, "Peer assessment in the digital age: a meta-analysis comparing peer and teacher ratings," *Assess Eval High Educ*, vol. 41, no. 2, pp. 245–264, 2016, doi: 10.1080/02602938.2014.999746.
- [6] D. Wiliam, "What is assessment for learning?" *Studies in Educational Evaluation*, vol. 37, no. 1, pp. 3–14, 2011, doi: <https://doi.org/10.1016/j.stueduc.2011.03.001>.
- [7] C. M. Brooks and J. L. Ammons, "Free Riding in Group Projects and the Effects of Timing, Frequency, and Specificity of Criteria in Peer Assessments," *Journal of Education for Business*, vol. 78, no. 5, pp. 268–272, 2003, doi: 10.1080/08832320309598613.
- [8] P. Aggarwal and C. L. O'Brien, "Social loafing on group projects: Structural antecedents and effect on student satisfaction," *Journal of Marketing Education*, vol. 30, no. 3, pp. 255–264, 2008, doi: 10.1177/0273475308322283.
- [9] D. Hall and S. Buzwell, "The problem of free-riding in group projects: Looking beyond social loafing as reason for non-contribution," *Active Learning in Higher Education*, vol. 14, no. 1, pp. 37–49, 2013, doi: 10.1177/1469787412467123.
- [10] E. Klein, "Peer Grading Tool - PGT," Website. Accessed: Apr. 15, 2024. [Online]. Available: <http://www.bfh.ch/peer-grading-tool>
- [11] O. Novogen, B. Nina, and E. A. Railean, "Assessment as Learning," *Assessment, Testing, and Measurement Strategies in Global Higher Education*, pp. 238–255, 2020, [Online]. Available: <https://api.semanticscholar.org/CorpusID:213527226>
- [12] R. Dann, "Assessment as learning: blurring the boundaries of assessment and learning for theory, policy and practice," *Assess Educ*, vol. 21, pp. 149–166, 2014, [Online]. Available: <https://api.semanticscholar.org/CorpusID:62881136>
- [13] S. Rahman and A. A. Majumder, "Is it assessment of learning or assessment for learning," *Southeast Asia Journal of Public Health*, vol. 4, pp. 72–74, 2015, [Online]. Available: <https://api.semanticscholar.org/CorpusID:73149255>
- [14] S. Chappuis and R. J. Stiggins, "Classroom Assessment for Learning.," *Educational Leadership*, vol. 60, pp. 40–43, 2002, [Online]. Available: <https://api.semanticscholar.org/CorpusID:142695229>
- [15] A. Loignon *et al.*, "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, 2017.
- [16] S. Stolz, "Peer-Grading Tool." Accessed: Apr. 15, 2024. [Online]. Available: <https://gitlab.com/st.stolz/Peer-Grading-Tool>
- [17] P. Sadler and E. Good, "The Impact of Self- and Peer-Grading on Student Learning," *Educational Assessment*, vol. 11, no. 1, pp. 1–31, Feb. 2006, doi: 10.1207/s15326977ea1101_1.
- [18] D. Nicol, A. Thomson, and C. Breslin, "Rethinking feedback practices in higher education: a peer review perspective," *Assess Eval High Educ*, vol. 39, no. 1, pp. 102–122, Jan. 2014, doi: 10.1080/02602938.2013.795518.
- [19] R. Albanese and D. D. Van Fleet, "Rational Behavior in Groups: The Free-Riding Tendency," *Academy of Management Review*, vol. 10, no. 2, pp. 244–255, 1985, doi: 10.5465/amr.1985.4278118.
- [20] E. Klein and C. Noppeney, "DigiPeG – Digitales Peer Grading für lernfördernde Gruppenarbeiten," Website. Accessed: Apr. 15, 2024. [Online]. Available: <https://belearn.swiss/en/projekt/digipeg-digitales-peer-grading-fuer-lernfoerdernde-gruppenarbeiten/>