How Students Manage Peer Feedback Through a Collaborative Activity in a CS1 Course

Géraldine Brieven[1], Laurent Leduc[2], Benoit Donnet[1]
[1]Institut Montefiore, Université de Liège, Belgium, [2]IFRES, Université de Liège, Belgium.

Abstract

In order to boost students' motivation in practicing their problem-solving skills and give them opportunities to get feedback, we broke our CS1 course routine with a disruptive cross-skilling activity. It relies on collaboration between teams of students where peer feedback (using rubric) stands as the cornerstone to design and build a solution responding to a given problem.

This paper aims at formally assessing the peer feedback process across three activity sessions. It also highlights the different success factors supporting peer feedback in that context through a cause and effect diagram. We show that peer feedback fosters primary problem-solving foundations. We also discuss its limitations, namely due to an insufficient granularity in the provided checklist as well as a lack of transversal skills from students, making them less comfortable with peer feedback. Although, by repeating the activity, students could manage it better and better and take more advantage of peer feedback.

Keywords: *Peer Feedback ; Team-based Learning ; Checklist ; Problem Solving ; CS1.*

1. Introduction

In a Introduction to Programming Course (CS1) dedicated to First Year students, a particular importance is given in teaching a problem-solving method as the resulting skills are crucial in the long run to be able to handle new challenges (Choudhar et al., 2022). Although, in practice, students often lack motivation during traditional programming exercise sessions, taking lots of time to solve a given exercise and missing feedback (Sharmin et al., 2020).

To overcome that issue, the Collaborative Design and Build (CDB) activity (Brieven et al., 2022) was deployed in our CS1 course. That activity implements Assembly Line Learning (Rosario et al., 2020) and Team-Based Learning (Burgess et al., 2021) where peer feedback is drawn on a checklist. To motivate that process, CDB was designed such as the reviewers need to give a feedback as clear and as precise as possible since they will have to rely on the productions they have reviewed to progress the solution. Previous work (Brieven et al., 2022) has shown that students feel motivated in taking part to CDB, due to its social dimension and its authentic aspect. CDB also appeared to boost students'enhancement in solving problems.

This paper addresses the particular focus of peer feedback, aiming to assess it and identify how it could be optimized. This is tackled through two research questions: (*RQ1*) How relevant is peer feedback in the context of CDB in a CS1 course ? (*RQ2*) Which prerequisites influence the feedback process in that context ? Answering those questions, this paper shows that students provide correct feedback based on a given checklist (RQ1) while their capacity to integrate it is moderate (RQ1). To explain that gap, we namely point out a lack of accuracy of the checklist criteria and students' difficulty in communicating. This diagnosis is corroborated through a *cause and effect diagram* depicting the feedback process in CDB in relation with other dimensions emerging from the discussion related to RQ1, such as students' skills (not only disciplinary) or the activity parameters (like the checklist) (RQ2).

2. CDB Activity

The CDB activity (Brieven et al., 2022) is made up of two phases: the *Design* (solution design) and the *Building* phase (solution implementation). Fig. 1 shows how CDB is setup.

The right side of Fig. 1 ("Classroom Configuration") illustrates that, for N participants, the CDB activity relies on G *Groups* of students, each Group being split in T *Teams* (each Team comprising S students, with $S \geq 2$). The goal of each Group is to solve T problems in a limited amount of time. The left part of Fig. 1 draws how the T problems are getting progressively solved, in parallel, over time, following the T steps required to frame a problem-solving process. This conception is inspired by real professional life as, in large-scale development projects, the different steps in solving a problem are performed by different teams.

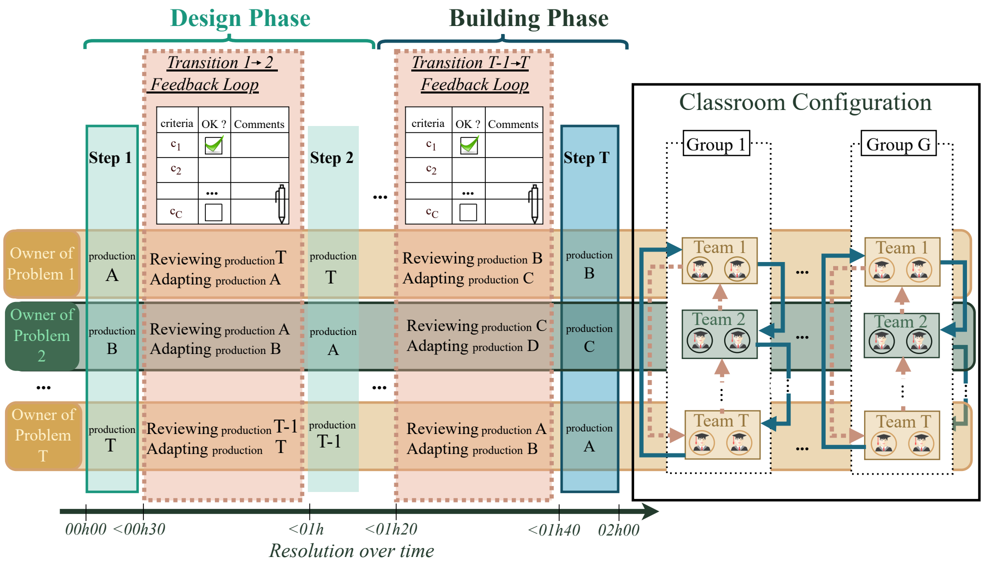

Figure 1. CDB generic set-up in the particular instance of two students per Team (S=2).

Going into more detail, at the beginning, each Team receives and gets responsible for one problem. For a given problem, the *T* steps are sequentially addressed, each Team being busy, turn-by-turn, with a specific one. At the end of each step, each Team work moves to the next Team, clockwise, similarly in each Group, as depicted by the plain arrows (Fig. 1, right part).

Then, a transition period is dedicated to allow each Team (*reviewers*) to report a feedback about the production provided by the previous Team (*submitters*). During that period, every Team holds two roles in parallel. For instance, Team 2 is reviewer of Team 1 (as expressed through the dashed arrow (Fig. 1, right part) linking the reviewers to the submitters) while it is submitter of Team 3. Note that the motivation behind that feedback loop is to limit the impact of a "poor quality work" on the next productions that are based on the previous ones.

Criteria supporting step 1	**OK ?**	**Comments (referred by tags if needed)**
Input(s)	☐	
Output	☐	
Objects that are used :		
- have a relevant name	☐	
- have a type that exists in C	☐	
- have a purpose that is clearly expressed	☐	
SP ...... :		
- is generic (i.e., doesn't restrict to an example)	☐	
- has a clear description	☐	
- has a functional description	☐	

Figure 2. Checklist related to the first step.

Session	*N*	*G*	*T*	*S*
1	63	10	3	2 or 3
2	49	8	3	2 or 3
3	48	8	3	2

Table 1. CDB parameters values.

The feedback is based on a rubric checklist (Bharuthram & Patel, 2017), as shown in Fig. 2. The first column lists the criteria a step output should meet. They are picked from the rubric supporting the course evaluations. In CDB, they are kept quite general, as the purpose is making students responsible for putting forward a solution rather than gradually disclosing it through the criteria. The next two columns allow each Team to specify whether a criterion is checked or not and attach comments to it. Once filled, the checklist is returned to the submitters who should adapt their work based on the boxes and comments in the checklist.

3. Method

Over the semester, three CDB sessions were organized, with an increasing complexity of problems to solve from one session to another. It is worth mentioning that participation was not mandatory. Table 1 summarizes the different paremeters values for each CDB session.

We collected data during those three sessions. In particular, on the one hand, an anonymous survey was addressed to students at the end of each session. Every survey included Likert scale questions, related to different aspects of the activity. On the other hand, all students' productions and feedbacks were collected and analysed afterwards. More specifically, first, each step production was qualified as reliable enough or not to support further the solution.

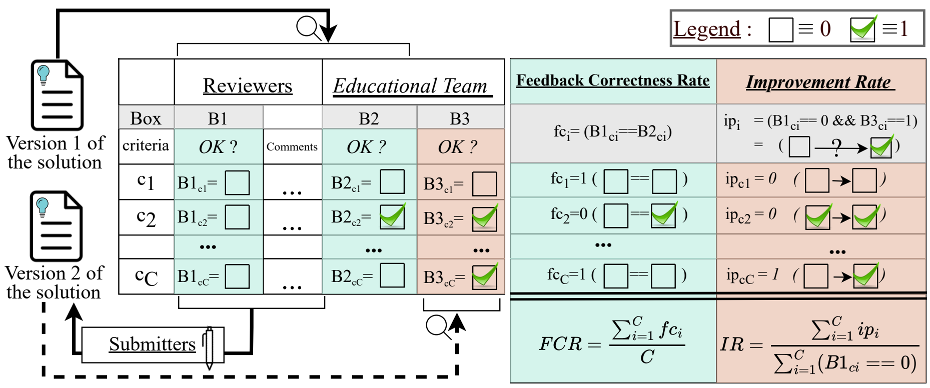

Figure 3: Illustration on how feedbacks were evaluated by the education team.

Next, each reviewers' checklist (reflecting the feedback) (see Fig. 2) was assessed. Fig. 3 depicts that process through a matrix in which the first column lists the *C* criteria related to a given step (each one being labeled c_i). The next column contains the boxes selected by the reviewers. Those are compared to the ones checked by the teacher (referred by the column "B2") used as benchmark to derive the *Feedback Correctness Rate* (FCR). Finally, the last column represents the checklist filled from the refined solution. It allows to compute the *Improvement rate* (IR), measuring how much the reviewers' feedback conducted submitters in meeting the unchecked criteria. For both rates, closer to 1 the better, closer to 0 the worst.

4. Results and Discussion

Fig. 4 exposes a cause and effect diagram where peer feedback is evaluated in CDB (RQ1). Then, the factors a successful feedback loop relies on are distilled (based on inferences emerging from RQ1 and the surveys' outcomes), classified, measured (on the grounds of students' views, at each session), and put in relation with each others (RQ2). Those links are quantified through the Pearson coefficient and its corresponding p-value.

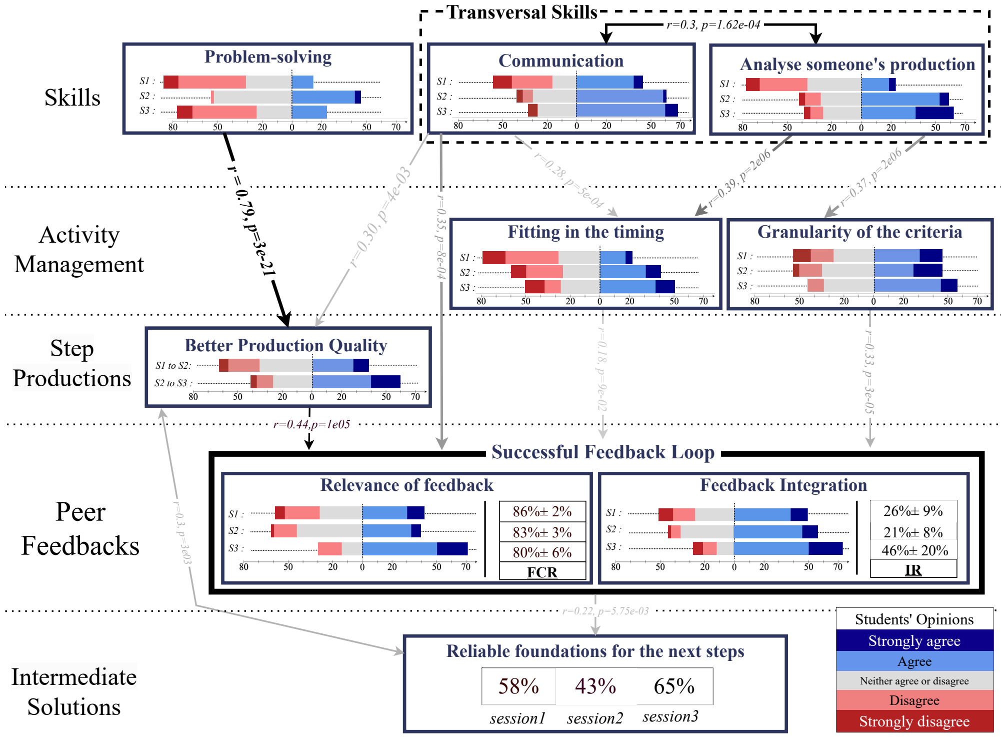

Figure 4. Cause and effect diagram depicting the feedback process in CDB.

4.1 (RQ1) How relevant is peer feedback in the context of CDB in a CS1 course ?

Considering the main frame of Fig. 4 ("Peer Feedbacks"), first, it shows that students seem quite mitigated about the relevance of the feedbacks they received during the first sessions. However, from the second to the last one, 30% of students found it richer, which is consistent to how comfortable they felt in integrating the feedback, rising from 49% to 72%. To further corroborate students' perception, it was put in perspective with results derived from their own productions and feedbacks. There, in contrast to students' opinion, a constant high FCR was computed. That overestimation with respect to students' opinion may be due to the fact that students missed tags in their production and comments as more accurate guidance. Those two feedback components were not involved in the FCR computation while they were taken into account in students' view. This explanation gets confirmed by noting that, over all the filled checklists, only 10% of criteria had a comment on the side. That may have prevented students from clearly identifying and fixing their gaps, especially those related to more complex criteria. Finally, the distribution of students behind each FCR mean is provided in Fig. 5a.

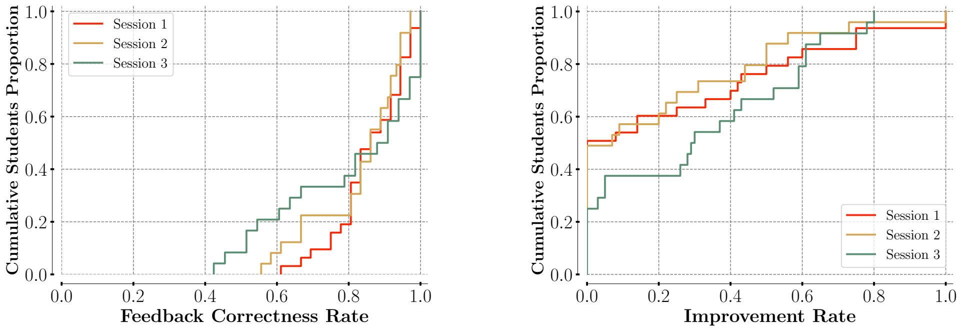

Figure 5. Cumulative distribution of students (a) emitting a correct feedback and (b) improving from a given feedback

It namely shows that, for the three sessions, 60% of reviewers checked at least 80% of the criteria correctly. It substantiates the gap between how submitters perceive the feedback and the actual feedback correctness in terms of checked boxes. In response to that, the list of criteria could be refined to tight better the expected solution (although it should not disclose it) and the reviewers should develop their feedback further too (through more comments typically). Next to this, Fig. 3 illustrates that the average IR remained limited (less than 30%) across the two first sessions but got better in the last one (reaching 46% (± 20%)). However, the large confidence interval related to that last result reflects the heterogeneity of submitters' perception in adapting to the feedback. That heterogeneity is distilled through Fig. 6b depicting that 25% of students did not improve at all, 35% improved on 50% to 80% of the criteria that were not inially met and the rest of the submitters standed in between. More generally, considering the three sessions, it can be noticed that submitters are likely to enhance from the feedback, but remaining gaps are persisting. Possibly, some students were missing time in tailoring their solution. Those results get consolidated seing the proportions of reliable refreshed production being above the average. However, they do not reach top values, which confirms that further refinements are still needed and suggests that students may lack problem-solving skills. It also converges to previous findings stating that feedback intervention shifts the attention away from the task itself and closer to the self (referring to, e.g., motivation, critical-thinking, communication) (Kluger and DeNisi (1996)), which is also partially highlighted through RQ2. From an evolution point of view, the proportion of reliable productions rises from 58% to 65% over the sessions despite the last session proposed harder statements, which emphases the production quality improvement over the CDB sessions.

4.2 (RQ2) Which prerequisites influence the feedback process in the context of CDB ?

Let us now strengthen the inferences stated in Sec. 5.1 by studying the factors assumed as influencing peer feedback. They are grouped in categories defined upon "Peer Feedbacks" in Fig. 5, and ordered based on how they impact on each others. Typically, students' "Skills" influence the way students handle the activity parameters (included in "Activity Management"). Then, student's ability to build and digest a feedback depends on the quality of the submission as well as how comfortable they are in managing the activity and mobilizing their skills.

In practice, all the components belonging to the three upper floors follow a similar trend than the feedback dimensions, getting better and better managed by students across the sessions, except the problem-solving skills. It is likely due to the nature of the problems that had to be tackled with a more formal angle students were not used to yet in the last session. Although, thanks to better transversal skills (Borova et al., 2021), students could counterbalance that difficulty by dealing better with the activity itself. More precisely, first, 80% of the students were initially struggling in fitting to the timing and, eventually, 50% were still facing that issue. That last proportion partially explains why the IR remains under the average : likely, besides some misunderstanding of the feedback submitters received, they lacked of time to analyze and integrate it. Next, the criteria of the checklist sounded missing accuracy for about half of the submitters (although their perception got more moderate), which comforts the idea that criteria should be refined. Finally, prior to that, students improved in collaborating with each others by communicating better and feeling more comfortable in analyzing the submitters' production, allowing them to deal better with the activity itself and, eventually, manage peer feedback, as underpinned through the positive correlation between those factors.

5. Perspective and Conclusion

To wrap up, CDB implements peer feedback embedded in a real-life scenario where students get responsible for sequential tasks on successive problems. To solve them, they need to give their best contributions,by providing valuable feedback and by shaping their outputs using feedbacks they receive in order to consolidate their ground and the ones of their reviewers.

In practice, in the context of CDB, students are quite good at identifying gaps in submitters' productions by relying on a checklist. However, digesting a feedback from reviewers appears more difficult (as also raised and deepened in other studies (Carless & Boud, 2018)). This paper explains that gap through three main reasons. First, students have difficulties in managing the limited time allocated to peer feedback. To overcome that, one could split the feedback loop period in two subparts to make sure students do not spend the whole period in building the feedback. Next, some criteria of the checklist lack of accuracy. Finally, most of the reviewers do not spontaneously distill them through further explanation, they only check boxes. In response to that, on the one hand, some criteria could be detailed and, one the other hand, students could improve their transversal skills that play a key role in peer feedback at both sides. Over the three CDB sessions, students felt more comfortable with respect to those last factors, which got reflected on the feedback process itself and the final productions quality. It confirms that CDB stands as an essential building block in our teaching methods, giving to students the opportunity to train transversal skills and learn from peer feedback.

References

Bharuthram, S., and Patel, M. (2017). Co-Constructing a Rubric Checklist with First Year University Students: A Self-Assessment Tool. *Apples - Journal of Applied Language Studies*, 11(4), 35–55. doi: 10.17011/apples/urn.201708073430.

Borova, T., Petrenko, V., Ved, T. and Pyvovarov, V. (2021). Developing Students' Transversal Competences Through Cultivation of Health Literacy Competence. *Marketing of Scientific and Research Organizations*. 40. 35-52. doi: 10.2478/minib-2021-0008.

Brieven, G., Leduc, L., and Donnet, B. (2022). Collaborative Design and Build Activity in a CS1 Course: A Practical Experience Report. In *8th International Conference on Higher Education Advances*. Universitat Polytecnica de Valencia. doi:10.4995/HEAd22.2022.14386.

Burgess, A., Roberts, C., Lane, A.S., Haq, I., Kalman, E., Pappalardo, N., and Bleasel, J. (2021). Peer Review in Team-Based Learning: Influencing Feedback Literacy. *BMC Medical Education*, 21(1), 426. doi: 10.1186/s12909-021-02821-6

Carless, D., and D. Boud. (2018). The Development of Student Feedback Literacy: Enabling Uptake of Feedback. *Assessment & Evaluation in Higher Education* 43 (8): 1315–1325. doi:10.1080/02602938.2018.1463354.

Choudhar, S., Bi, N., Singh P. N., and Talwar P. (2022). Study on Problem Solving Skills and Its Importance. *World Journal of English Language,* 12(3), 1-47. doi: 10.5430/wjel.v12n3p47.

Harackiewicz, J. M., Smith, J. L., and Priniski, S. J. (2016). Interest Matters: The Importance of Promoting Interest in Education. *Policy Insights from the Behavioral and Brain Sciences*, 3(2), 220–227. doi: 10.1177/2372732216655542.

Kluger, A. and DeNisi, A.. (1996). The Effects of Feedback Interventions on Performance: A Historical Review, a Meta-Analysis, and a Preliminary Feedback Intervention Theory. *Psychological Bulletin*. 119. 254-284. doi: 10.1037/0033-2909.119.2.254.

Rosario, J., Lebowitz, D., Webb, A. L., Ganti, L., Vera, A., Macintosh, T., Walker, A., and Rubeor, J. (2020). Assembly Line Education: A Novel Educational Technique for Today's Learners. *Cureus*, 12(2). doi: 10.7759/cureus.7065.

Sharmin, S., Zingaro, D., and Brett, C. (2020). Weekly Open-Ended Exercises and Student Motivation in CS1. *20th Koli Calling International Conference on Computing Education Research*. ACM. doi: 10.1145/3428029.3428035.