

Flagging in teacher-facing orchestration dashboards: factors affecting its use in Pyramid CSCL debriefing

Max Dieckmann
TIDE Research Group
Universitat Pompeu Fabra (UPF)
Barcelona, Spain
0000-0001-7128-8337

Davinia Hernández-Leo
TIDE Research Group
Universitat Pompeu Fabra (UPF)
Barcelona, Spain
0000-0003-0548-7455

Ishari Amarasinghe
TIDE Research Group
Universitat Pompeu Fabra (UPF)
Barcelona, Spain
0000-0003-2960-4804

Abstract—Teacher-led debriefing has the potential to positively affect learning gains when conducted at the end of collaborative learning activities. In order for debriefing to be effective, the teacher needs to base it on the learner’s process during the activity. Research in the field of Computer-Supported Collaborative Learning (CSCL) is proposing teaching-facing dashboards as tools that facilitate the monitoring and orchestration of activities. However, research has paid less attention to how these dashboards can support debriefing. We explore how adding a “flagging” feature to a CSCL orchestration dashboard can support debriefing by reporting a qualitative preliminary study in which the flagging feature was used during a Pyramid CSCL script activity. Results indicate that the dashboard interface design, number of student responses, number of errors in student responses, and whether student responses meet the teachers expectations most influence the use and utility of the feature. Additionally, we identified avenues for improving and extending the design of the feature.

Index Terms—computer-supported collaborative learning, flagging, teacher-led debriefing

I. INTRODUCTION

In a Computer-Supported Collaborative Learning (CSCL) activity, students use computers to connect with each other and work together on a task, while the same technology helps the teacher take on the role of a conductor who orchestrates the activity: they provide scaffolds, help struggling students, and control the sequence of events [1], [9]. A CSCL activity can follow a script that provides a structure for the activity - defining its goals, sequence of events, rules for forming the student groups etc. Some CSCL scripts have been implemented as applications that the students access via mobile devices and / or computers. These applications provide the opportunity for teacher-facing dashboards, which enable the teacher to orchestrate the activity and provide them with log data and learning analytics on the students’ work [2].

A CSCL activity can be followed by a teacher-led debriefing session, where the activity is reflected on: misconceptions and errors are being revealed and discussed, the material is put into context, and feedback is given [3]. There are indications that the feedback and revelations of connections to real-life contexts that debriefing provides has many positive effects which ultimately lead to learning gains [4].

However, in order for debriefing to be successful, the teacher has to react to the students’ inputs “on the fly”, as the feedback has to address those issues that came up [3].

To meet this demand and support debriefing, several options need to be explored. Providing teachers with log data in real time has been shown to be beneficial for orchestration in a CSCL context [5], [8]. We conducted an exploratory study where we examined how log data can facilitate debriefing by introducing a “flagging” feature to the teacher-facing dashboard of a CSCL application. This feature allows for the user to flag (mark) incoming student answers as particularly relevant for later discussion. These flagged answers are then displayed in the dashboard interface for later reference.

In this paper we attempted to answer the research question “What factors influence the utility and utilisation of a feature that allows teachers to flag student answers for later reference in a CSCL activity dashboard?”

Section 2 of this paper describes the method we used to conduct our study. Results are presented in Section 3. In Section 4, we discuss the results and outline directions for future work.

II. METHOD

We conducted an exploratory qualitative study, where four teachers ran CSCL activities with undergraduate students (see Table I for more details). In total, we collected data from 11 sessions. The class size ranged from 40 - 80 students ($\mu = 60.3$, $\sigma = 12.7$). Teachers were asked to use the flagging feature in eight of these sessions, the feature was intentionally unused in the remaining three. The activity followed the “pyramid” script and the teachers used “PyramidApp”, a software that implements a special version of the Pyramid script and includes a teacher-facing dashboard [6]. The activities were each followed by a debriefing session where the teacher went over the student responses and addressed key issues.

The “pyramid” script runs in phases [7]: In the first phase, the students each individually respond to the task that the teacher selected for the activity. In the second phase, they are automatically assigned to groups and then anonymously rate the answers their group members gave. In the third phase, they collaborate with their group members to formulate an answer the whole group agrees with. For this, they have access

TABLE I
DETAILS ON THE CSCL ACTIVITY SESSIONS

Activity question	Teacher	Sessions
What is considered plagiarism and why is it important to know it in university?	A	3
Read the given case study carefully and identify two ethical principles that have been applied in the particular case, and one that has not been applied. Briefly justify your selection.	B	3
What surprised you most about what you learned in the most recent lectures? Why?	C	2
Using the provided data, define a user profile oriented towards financial risk. Choose the 4 characteristics you believe are most important, explain how you would generate them and argue why they are related to financial risk.	D	3

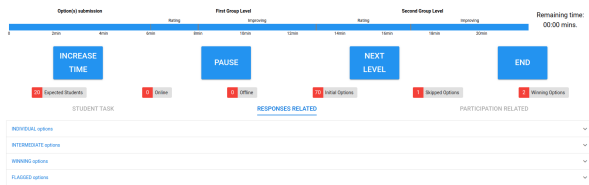


Fig. 1. The “PyramidApp” dashboard. Remaining time is displayed in the top right corner, responses in the centre. The buttons expand lists for individual responses, intermediate, winning and flagged responses, respectively.

to a text chat and a collaborative text editor, as well as the individual answers and their respective ratings. Depending on the class size, phases two and three are repeated with increasingly bigger groups, such that in the end only a few answers remain that represent the group consensus.

The current version of “PyramidApp” gave teachers the opportunity to flag incoming individual student answers that were generated in phase one for later reference (e.g. during debriefing). This allows teachers to keep track of competing ideas, unusual solutions and common errors that might get lost in the later stages when group answers are formulated. These answers were then displayed in a separate tab in the interface (see Figure 1).

Teachers were given a questionnaire after completing each session, where they reflected and commented on the activity. A researcher present during the experiment took observation notes on the behaviour of the teacher and screen recordings of the computer with the teacher-facing dashboard were taken.

III. RESULTS

We analysed the observation notes to determine at what times the teacher was focusing on the dashboard and by studying the mouse movements in the screen recordings, we were able to discern that some teachers flagged student answers but never expanded the tab where these answers were then placed in the interface. All teachers did however study the final answers at the end of the activity before initiating the debriefing. This was also evident in the questionnaires, where

all teachers mentioned seeing the final answers as one of the most helpful features of the dashboard. Other features that were praised were the display of incoming student answers, remaining time and status alerts.

A further examination of the questionnaires revealed that one teacher (teacher A) noticed that being told to use the flag feature made them focus more on reading incoming student answers, but that another reason for focusing less during the no-flagging condition was due to already being familiar with the kind of answers the students gave (all sessions were on the same day and the no-flagging session was last). This coincides with our observations that they spent more time reading student responses the first time they ran the activity than the second time they ran the same activity with a different cohort of students. Teacher A noted that looking at the flagged options helped them directly to prepare the debriefing and that they remembered the important points in student answers because they were paying attention to the student answers earlier due to the flagging. Teacher A also suggested that the feature would be most useful when there is a lot of confusion in the student answers that need to be addressed. According to the teachers, this was not as much the case in these sessions, which is likely related to the relative simplicity of the questions the students had to answer (see Table I).

Another teacher (teacher B) mentioned not being aware of the flagging feature in the first session and even after using it in the second not fully understanding its functionality. Teacher B suggested that a tutorial video could help explain the dashboard to the teacher and make it easier to use.

Additionally, several weaknesses in the dashboard interface design were evident through observations or made explicit in teachers’ comments: Incoming student answers were placed in a list, which would automatically scroll with new entries. At times, this made it difficult to keep the focus on a specific answer the teacher was trying to read. The placement of flagged answers separately to the final answers was questioned, in particular as some teachers did not appear to search them out before starting the debriefing. The teachers also suggested that pre-processing and displaying the student answers differently could make it easier. Some ideas that were mentioned were grouping similar answers, word clouds, and a visualisation of how students’ answers change over the course of the activity.

IV. DISCUSSION

We were able to identify several factors that influence the use of the flagging feature (summarised in Table II): first and foremost, the teachers have to be aware of the existence and objective of the feature. This depends on the structure of the dashboard interface, as well as the placement of the elements that form part of the feature. Guiding elements such as tutorial videos or popup hints could influence this factor. Teachers need to be made aware where flagged answers are placed, as they otherwise never refer to them again. Another factor seems to be that the feature is used less the more the teacher has repeated the same activity and is familiar with the (expected) student responses.

TABLE II
FACTORS INFLUENCING THE USE / UTILITY OF THE FLAGGING FEATURE.

Factor	Relevant teacher comment
<i>No. of student responses:</i> Flagging is particularly useful when dealing with a large number of responses, as they are often similar and can be condensed to a short list where the essential student ideas are represented.	“I also didn’t want to have many options flagged so as not to have an unmanageable long list.” (Teacher A) “We had many answers to the same question [...]” (Teacher A)
<i>No. of errors in student responses:</i> A very important part of debriefing is addressing student misunderstandings. Thus, wrong answers are great candidates for flagging.	“[...] I think that in other types of tasks (where more mistakes/confusion may occur), I would have used more the dashboard [sic] info for debriefing.” (Teacher A)
<i>Content of student responses:</i> Flagging becomes most useful when the teacher notices something interesting or surprising in a student answer that they want to comment on, but even having an example of a model answer can be useful for debriefing.	“However, it was good to know that I had a list of ideas to consider in the flag and winning lists - just in case.” (Teacher A) “But I do wanted [sic] to keep several interesting options flagged to mention at the end.” (Teacher A)
<i>Awareness of the feature and its intended use:</i> Teachers need to be aware of the feature and understand why and how they can use it to help prepare the debriefing session.	“For instance I was not aware of the Flag option [...], and still now I am not fully sure about its functionality.” (Teacher B)
<i>Dashboard interface design / User experience:</i> A clear and bug-free interface that makes using the feature intuitive for the user will make it more likely for teachers to adopt it.	“when receiving answers the new ones coming in move down the previous ones and it’s difficult to read them” (Teacher C) “Show the flagged responses together with the final response” (Teacher D)
<i>Teacher familiarity with the activity and common errors:</i> When repeating the same activity with new students, most errors and types of student responses are already known. Thus, there is less need to focus on and flag the incoming student answers.	“[...] I was not reading much what students [sic] were writing. Reasons were: 1) I was not flagging so I was more relaxed in reading, 2) as it was the same third task of the day I already had a good overview of students types of answers (everything was being repeated, including the only one aspect I identified that was needing clarification in the debriefing).” (Teacher A)

We also found a weakness in the dashboard interface that influenced the usability of the feature — namely that flagging answers was at times unnecessarily difficult due to the list of answers automatically scrolling. As mentioned before, debriefing needs to be conducted “on-the-fly”, in reaction to the students’ responses [3]. It is therefore not surprising that the teachers see the flagging feature as most useful when the students’ answers are surprising to them or when there are a lot of responses to go through. As such, the familiarity of the teachers with the material and its pitfalls, the number of student responses (i.e. the class size), the variety of responses, and the number of mistakes and confusions in the responses are all relevant.

The study was exploratory with only a handful of participat-

ing teachers and thus had a focus on identifying promising directions for future work. In addition to the factors we identified that can be used to inform developments to extend and improve the flagging feature, we can turn to the suggestions the teachers made: First, the student responses could be automatically pre-selected or grouped by their content. This has the potential to alleviate some of the problems associated with a large number of responses and make it easier for the teacher to identify the key issues that they would like to address [10]. Secondly, visualisations could be implemented that help the teacher get information on the state of the activity at a glance. This could take the form of word clouds, graphics showing the changes in answers over time, or group indicators for low activity / agreement. The last suggestion was to extend the dashboard with a debriefing notebook, into which (parts of) student answers could be copied-and-pasted and the teacher could add thoughts and comments. This could make the intended use of flagging student answers clearer, as well as allow for a more flexible debriefing preparation.

V. ACKNOWLEDGMENTS

This work has been partially funded by the Ministry of Science and Innovation and the National Research Agency (PID2020-112584RB-C33/MICIN/AEI/10.13039/501100011033) and ICREA under the ICREA Academia programme (D. Hernández-Leo, Serra Hunter).

REFERENCES

- [1] Dillenbourg, Pierre, Sanna Järvelä, and Frank Fischer. “The evolution of research on computer-supported collaborative learning.” In *Technology-enhanced learning*, pp. 3-19. Springer, Dordrecht, 2009.
- [2] Dillenbourg, Pierre, and Patrick Jermann. “Designing integrative scripts.” In *Scripting computer-supported collaborative learning*, pp. 275-301. Springer, Boston, MA, 2007.
- [3] Dillenbourg, Pierre. “Integrating technologies into educational ecosystems.” *Distance Education* 29, no. 2 (2008): 127-140.
- [4] Hattie, John. *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. Routledge, 2008.
- [5] Matuk, Camillia, Michael Tissenbaum, and Bertrand Schneider. “Real-time orchestrational technologies in computer-supported collaborative learning: an introduction to the special issue.” *International Journal of Computer-Supported Collaborative Learning* 14, no. 3 (2019): 251-260.
- [6] Manathunga, Kalpani, and Davinia Hernández-Leo. “Authoring and enactment of mobile pyramid-based collaborative learning activities.” *British Journal of Educational Technology* 49, no. 2 (2018): 262-275.
- [7] Hernández-Leo, Davinia, Eloy D. Villascclaras-Fernández, Juan I. Asensio-Pérez, Yannis Dimitriadis, Iván M. Jorrín-Abellán, Inés Ruiz-Requies, and Bartolomé Rubia-Avi. “COLLAGE: A collaborative Learning Design editor based on patterns.” *Journal Of Educational Technology And Society* 9, no. 1 (2006): 58.
- [8] Amarasinghe, Ishari, Davinia Hernández-Leo, Konstantinos Michos, and Milica Vujovic. “An Actionable Orchestration Dashboard to Enhance Collaboration in the Classroom.” *IEEE Transactions on Learning Technologies* 13, no. 4 (2020): 662-675.
- [9] Dillenbourg, Pierre, and Patrick Jermann. “Technology for classroom orchestration.” In *New science of learning*, pp. 525-552. Springer, New York, NY, 2010.
- [10] Alvarez, Claudio, Gustavo Zurita, Andrés Carvallo, Pablo Ramírez, Eugenio Bravo, and Nelson Baloian. “Automatic Content Analysis of Student Moral Discourse in a Collaborative Learning Activity.” In *International Conference on Collaboration Technologies and Social Computing*, pp. 3-19. Springer, Cham, 2021.