13th LYON CSCL2019

A Wide Lens: Combining Embodied, Enactive, Extended, and Embedded Learning in Collaborative Settings:

13th International Conference on Computer Supported Collaborative Learning

Conference Proceedings Volume 1

Edited by
Kristine Lund, Gerald Niccolai, Elise Lavoué, Cindy Hmelo-Silver
Gahgene Gweon, Michael Baker

17 - 21 June 2019
École Normale Supérieure de Lyon, France

ENS DE LYON

ISLS

A Wide Lens: Combining Embodied, Enactive, Extended, and Embedded Learning in Collaborative Settings

13th International Conference on Computer Supported Collaborative Learning (CSCL) 2019

Volume 1

13th International Conference on Computer Supported Collaborative Learning

June 17-21, 2019, Lyon, France
École Normale Supérieure de Lyon, France

Editors:
Kristine Lund, Gerald Niccolai, Elise Lavoué, Cindy Hmelo-Silver
Gahgene Gweon, and Michael Baker

A Wide Lens: Combining Embodied, Enactive, Extended, and Embedded Learning in Collaborative Settings, 13th International Conference on Computer Supported Collaborative Learning (CSCL) 2019
June 17-21, 2019, Lyon, France
École Normale Supérieure de Lyon, France.

www.isls.org

ISBN: 978-1-7324672-3-1 (Volume 1, PDF Version)
ISSN: 1573-4552

Cite as: Lund, K., Niccolai, G., Lavoué, E., Hmelo-Silver, C., Gweon, G., and Baker, M. (Eds.). (2019). *A Wide Lens: Combining Embodied, Enactive, Extended, and Embedded Learning in Collaborative Settings, 13th International Conference on Computer Supported Collaborative Learning (CSCL) 2019, Volume 1.* Lyon, France: International Society of the Learning Sciences.

Adaptive Support for Collaboration on Tabletop Computers

Abigail C. Evans, Khoury College of Computer Sciences, Northeastern University, ab.evans@northeastern.edu
Katie Davis, Information School, University of Washington, kdavis78@uw.edu
Jacob O. Wobbrock, Information School, University of Washington, wobbrock@uw.edu

Abstract: We present the design, implementation, and evaluation of a system providing adaptive support for collaborative learning at a tabletop computer. Adapting to support collaboration involves tackling several sub-problems: detecting when a group is struggling, determining when the system should provide support, what sort of support it should provide, and for how long. A classroom evaluation of our system showed it was effective at detecting collaboration breakdowns and provided preliminary evidence that just-in-time adaptive support for collaboration might help to deter disruptive behavior in small group settings.

Introduction

Computer-supported collaborative learning (CSCL) environments are becoming increasingly sophisticated, enabling new ways for students to work together. However, prior research has demonstrated that students do not always know how to collaborate effectively, which can inhibit the success of small group learning, e.g., (Dillenbourg & Jermann, 2007; Järvelä & Hadwin, 2013; Rogat & Linnenbrink-Garcia, 2011). The findings of this prior work suggest that collaboration itself is a skill that needs to be developed in the classroom.

In this work, we focus on face-to-face, small group collaborative learning at tabletop computers because their large, shared touch screen makes them well-suited for collaboration. However, the mere presence of a large touch screen cannot make up for a lack of collaboration skills—students who do not know how to collaborate effectively in traditional face-to-face settings will continue to struggle at a tabletop computer. Our long-term goal has been to enable tabletop computers to make up for poor collaboration skills. Existing work has addressed the first step in realizing this goal, namely detecting collaboration breakdowns (Evans, Wobbrock, & Davis, 2016; Evans & Wobbrock, 2014). In this paper, we present our work toward the next step—designing ways for the tabletop interface to adapt when breakdowns are detected in order to encourage more effective collaboration. We describe the design and implementation of our system, which was guided by learning theory and educational research, and a classroom evaluation. Our analysis focuses on how the configuration of the adaptations appeared to influence group dynamics over time and through repeated exposure for one group of students particularly prone to collaboration problems. Our key findings are that the adaptations deterred disruptive behavior and halved the length of periods of low-quality collaboration for this group.

Measuring collaborative learning

To determine what qualifies as *effective* collaborative learning, we use the concept of social regulation, which refers to "the social processes groups use to regulate their joint work on a task" (Rogat & Linnenbrink-Garcia, 2011). Social regulation is an extension of the concept of self-regulation in individual learning to groups of learners in a collaborative setting (Järvelä & Hadwin, 2013; Panadero & Järvelä, 2015; Rogat & Linnenbrink-Garcia, 2011; Volet, Vauras, & Salonen, 2009). Social regulation typically refers to metacognitive processes independent of domain knowledge, although this relationship needs further research.

Rogat and Linnenbrink-Garcia (2011) identify three dimensions of social regulation: *planning* the group's approach to a task, *monitoring* of understanding and progress, and *behavioral engagement*—efforts to get group members to engage with the task. A group's use of each dimension can vary in quality, with high-quality social regulation processes leading to socially-shared regulation, in which all group members maintain joint attention on the learning activity, regularly checking in on the goals, plans, and progress. Low-quality social regulation processes, such as when groups fail to come up with an appropriate plan, can lead to poor learning outcomes. We chose to follow Rogat & Linnenbrink-Garcia's framework when coding for social regulation. Their approach was selected above others because of their detailed descriptions of what high-and low-quality social regulation processes look like in practice; thus, it could easily be applied to our own work.

Adaptive support for collaborative learning

There has been extensive prior work on adaptive support for collaboration in text-based CSCL environments. For example, guiding systems coach students through a collaboration using adaptive feedback, e.g., (Hadwin, Oshige, Gress, & Winne, 2010; Kumar & Rosé, 2011; Wang, Rosé, & Chang, 2011), and group awareness tools provide feedback to students, often in the form of visualizations, on how they are collaborating, e.g., (Järvelä et al., 2014; Malmberg, Järvelä, Järvenoja, & Panadero, 2015; Trausan-Matu, Dascalu, & Rebedea, 2014).

Many of the group awareness tools described in the literature make use of the closed nature of text-based collaborative environments, in which all group interactions can be captured and mined for patterns associated with the quality of collaboration. In face-to-face group work at a tabletop, however, the computer can only access direct human-tabletop interaction—the verbal and gestural interactions that learners have with each other cannot be captured without the use of external sensors. Martinez-Maldonado et al. used external sensors to capture learners' verbal and physical interactions during tabletop collaboration in order to visualize each group member's level of participation (Martinez-maldonado, Kay, Yacef, Edbauer, & Dimitriadis, 2013; Martinez-Maldonado, Yacef, Kay, Al-Qaraghuli, & Kharrufa, 2011; Martinez, Collins, Kay, & Yacef, 2011). Although these visualizations were aimed at teachers, they could also be presented to students as group awareness tools. At the time of this writing, the work by Martinez-Maldonado et al. remains the only other research providing real-time adaptive support for collaborative learning at tabletop computers.

It is important to consider how adaptive support should be presented to students and situated in the larger context of a learning activity or course. Wise (2014) notes it is not enough to simply present students with interventions—if students are to benefit from an intervention, they need to understand its intent and how to engage with it. An intervention should be an "integral part of course activity tied to goals and expectations" (Wise, 2014, p. 206). Interventions should also encourage students to have agency in their own learning.

System design

Providing real-time adaptive support for collaboration involves two major components: (1) *detection* of behavior that will trigger adaptations, and (2) the software *adaptations* themselves. A third consideration is when and how adaptations are triggered, which we refer to as the *configuration*.

Detection

We used Evans et al.'s approach to detecting the quality of collaboration at a tabletop computer using a group's touch-interaction patterns (Evans et al., 2016). Evans et al. identified two touch patterns that serve as reliable indicators of a group's collaboration quality: *unrelated touches* (UT), and *overlapping unrelated sequences* (OUS). A *sequence* is a series of touches carried out by an individual that represents a complete action. UT measures the proportion of touches in a sequence that involve objects that are unrelated in the context of a learning activity, with a greater proportion of touches to unrelated objects corresponding to lower quality collaboration—specifically, off task behavior or failure to settle on a suitable plan to complete the activity. OUS is a measure of whether a single student or multiple students are interacting with the tabletop, and when multiple students are interacting, whether they are working with objects that are related to each other in the context of the activity. Evans et al. found that, during periods of high-quality collaboration, students either took turns to interact with the computer or multiple students worked together with related objects. Multiple students interacting with unrelated objects was a sign that students were either working independently or off-task.

Evans et al.'s detection approach was built in to our learning software following the implementation details described therein (Evans et al., 2016)—two-minute intervals of touch data are checked against the two touch patterns, UT and OUS, each of which returns a label estimating the quality of collaboration occurring in that two-minute period as either high-, medium-, or low-quality. It is important to note that the touch patterns need to be able to distinguish the touches of individual students from those of other group members, which most tabletop computers are unable to do. We also therefore implemented Evans et al.'s *Group Touch* method of distinguishing among tabletop users to meet this requirement (Evans, Davis, Fogarty, & Wobbrock, 2017).

Adaptations

We aimed to create adaptations that would be application-independent in principle. Therefore, we developed the adaptations without a specific application in mind. Initial brainstorming resulted in 11 adaptation ideas, of which four were implemented and pilot tested. See Figure 1, below, for screen shots of these adaptations:

The *group awareness icon* draws on prior work on group awareness tools e.g., (Järvelä et al., 2014; Malmberg, Järvelä, Järvenoja, & Panadero, 2015; Trausan-Matu, Dascalu, & Rebedea, 2014). An icon remains visible on screen for the duration of the activity and changes color according to the detected collaboration quality, with green representing high-quality collaboration, amber representing medium-quality collaboration, and red representing low-quality collaboration.

Control lockout blocks or disables select controls, such as buttons or menu items, to reduce the number of active application elements. The rationale for this adaptation comes from prior work that found that when group members jump around between many different controls, they are typically either off task or struggling to identify or stick to a task plan (Evans et al., 2016). The aim is to encourage sustained whole-group focus on a specific piece of the assigned activity.

Voting prevents actions that affect the global state of the application, such as deleting work or marking an activity complete and moving on, from being carried out unless the majority of group members vote to proceed. The goal is to encourage discussion and prevent individual students from dominating the group.

Escalate to authority sends an alert to the teacher's mobile device requesting human intervention.

Figure 1. The adaptations in the study applications. Top left: the group awareness icon, positioned in the corners of the screen, showing amber (medium-quality collaboration). Top right: the prompt, added after the pilot. Bottom left: control lockout has disabled a control that allows students to switch the task. Bottom right: voting blocks actions from being carried out until a majority of the group members vote for it.

Configuration

The adaptations aim to scaffold effective collaborative interactions by making it difficult for students to interact with the screen in ways deemed to be ineffective. However, no single adaptation covers the full range of desirable collaborative behaviors, all of which need to be adopted by the group if they are to work together effectively. Therefore, we decided that the adaptations should be triggered in a sequence, where each new adaptation is added to those already in use but is only triggered if the collaboration quality does not improve. Layering adaptations can make it easier for students to ground themselves in the activity's expectations, because together the adaptations give a more complete picture of how to be effective than any isolated adaptation. When the collaboration does improve, all active adaptations are removed at once—an example of "fading," which, in the scaffolding literature, refers to the process of removing supports that are no longer needed (Pea, 2004).

Removing adaptations when collaboration improves also supports Wise's (2014) principle of agency. After a group has experienced a sequence of adaptations once, they will know what is coming next if the sequence begins again. Because the students can anticipate the next step in the sequence, they also have the opportunity to preemptively adopt the behaviors that the next adaptation will enforce, removing the need for that adaptation and thereby preventing it from being triggered. For example, if students know that the next adaptation to be triggered if their collaboration does not improve will be *control lockout,* they can aim to improve by interacting only with objects directly relevant to the aspect of the task they are working on.

The sequence begins with the *group awareness icon*, which remains visible on-screen for the whole activity, changing color according to the detected collaboration quality. The second adaptation, *control lockout*, is triggered when the icon turns red, representing poor collaboration quality, and no other adaptations are active. *Control lockout* was chosen as the first restrictive adaptation based on observations from prior work on tabletop collaboration that groups often jumped into an activity without taking time to plan or check the instructions (Evans et al., 2016). If collaboration does not improve, the third adaptation, *voting*, is added. *Voting* can slow down progress and become frustrating, so it is only triggered after students have had the opportunity to reflect on and improve their collaboration. Finally, if collaboration still does not improve after a period of voting, the application will *escalate to authority* and call the teacher to intervene. There are many other ways to combine, order, and layer the selected adaptations in a sequence. This particular sequence was chosen initially as the adaptations progress in order from least to most restrictive or interventionist, giving students time to reflect on and change their collaboration behavior before increasing the level of restrictions imposed on the group.

In addition to deciding which adaptations to build and the order in which they should be triggered, we had to decide exactly when an adaptation should be triggered, how it should be presented and explained to students, how long each adaptation should remain active, and when adaptations should be removed. Each of these choices could impact the outcome of the intervention. The criteria we chose for triggering adaptations were derived from Evans et al.'s (2016) analysis of the reliability of using the touch patterns (UT and OUS) to detect collaboration quality. Adaptations were triggered immediately if both touch patterns labeled a single interval low quality, or if two consecutive intervals were labeled low or medium quality by UT. Adaptations remained active until both touch patterns produced high quality labels for four consecutive minutes.

Classroom evaluation

The prototype adaptations were evaluated in a classroom setting. Our research questions were: (1) Are the adaptations triggered appropriately? (2) How do students respond when an adaptation is triggered?

Participants and apparatus

Eleven 10^{th} and 11^{th} graders (9 male, 2 female) participated in the study. The students were enrolled in a six-week beginner's course on video game development offered as part of a college preparation program serving low income students. The course used in this study was one of the program's elective options. The course focus was educational games, and students were asked to design a game to raise awareness of an environmental issue: snow leopard conservation. In the first part of the course, the students learned about snow leopards and ongoing efforts to protect the species. In the second part of the course, the students learned the basics of video game development using Unity (a 3-D game development platform: see https://unity3d.com/). All but one of the students were new to writing code and all were new to video game development and Unity. In the third and final part of the course, the students worked on their game projects.

The students were randomly assigned to three small groups—two groups of four and one group of three. The students stayed in the same groups for all collaborative activities. The groups used three custom-built learning applications across four sessions at a Microsoft Surface Hub, which features a 55" multitouch screen. The applications addressed the learning objectives for the class sessions in which they were to be used and were designed to be used alongside other activities.

Snow Leopards 101 is an introduction to snow leopards and their ecosystem adapted from an existing non-tabletop curriculum (Facing the Future & Snow Leopard Trust, 2009). In *Help a Scientist*, students take on the role of scientists studying wild snow leopards in Mongolia, gathering and analyzing photographic data. *Game Challenge* was used during the third part of the course. This application featured a partially created game world that students could build upon and adjust to create different outcomes while learning important and complex concepts used in Unity, such as how to work with its built-in physics engine.

The Surface Hub is primarily intended to be used in a vertical orientation but for this study it was placed flat on a table in a small breakout room across a hallway from the main classroom. A wide-angle video camera was mounted on a wall so that students could freely move around the table. The camera was angled toward the screen so that it could capture every touch and the interactions among the group members. Ten group sessions at the tabletop were video recorded for this study and the computer logged every touch.

Study design

The study was split into three phases: (1) baseline data collection; (2) a pilot test of the initial implementation of the adaptations, which resulted in a revised implementation; and (3) an evaluation of the final implementation. The *Snow Leopards 101* application was used for baseline data collection, *Help a Scientist* was used for the pilot test, and *Game Challenge* was used for the final evaluation taking place two and a half weeks after the pilot. Students used each application for 25 minutes per session. *Game Challenge* was used over two sessions.

The adaptations were piloted with Group 1 only. The final adaptations were available for Group 1 and Group 3 during the final evaluation—if poor quality collaboration were detected, the adaptations would trigger. Group 2 served as a control so the adaptations were not available for them during phase 3, although the collaboration quality labels output by the touch patterns were still recorded. We wanted to look for differences in how groups dealt with collaboration problems with or without the intervention of adaptations. Students were randomly assigned to groups and groups were randomly assigned to conditions in an attempt to reduce differences between groups. However, given that it was unlikely the groups would be truly equal in a study of this scale, the goal of the baseline data collection was to understand the differences between groups.

The first time a group used the tabletop with adaptations enabled, a researcher explained to them that the computer was tracking how they interacted with it in order to help them work together effectively. They were told that the color of the *group awareness icon* updated every minute to give them feedback on how they were collaborating and that if it turned red or stayed amber for several minutes, the computer would block some controls or ask them to vote before carrying out certain actions. They were also told that the adaptations would go away if the icons stayed green for four minutes, and that they could keep the icons green by maintaining a shared focus on the assigned task, discussing the content, and listening to each other's ideas.

Pilot test

The adaptations were piloted with Group 1, a group of four boys randomly assigned to the pilot, as they used the *Help a Scientist* application. The *group awareness icon* (Figure 1, top left) was displayed in two of the corners of the screen so that it would be visible to all group members without obstructing the work area. *Control lockout* (Figure 1, bottom left) was configured to trigger if the icon turned red when no other adaptations were active. Before activation, a warning popped up on screen stating, "Some buttons and controls will be temporarily disabled to help you to stay focused on the task." The warning remained on screen for 30 seconds, blocking all interaction for the time that it was on screen. *Voting* (Figure 1, bottom right) was also preceded by a 30-second warning message: "You will temporarily be asked to vote in order to carry out certain actions such as changing activity or closing windows." If a group was asked to vote on an action, a message was shown describing what the group members were voting on. Finally, *escalate to authority* would send an alert to the teacher that the group might be struggling. Unlike *control lockout* and *voting*, *escalate to authority* was invisible to the students.

To evaluate the piloted adaptations, the first author reviewed the video of the pilot session and the computer's log files. Given that the purpose of the pilot was to get feedback on a number of design choices quickly enough to make changes in time for a formal evaluation two and a half weeks later, it was not possible to do a full in-depth analysis of the video at this point. Instead, observations were made of what the group was doing in the run up to an adaptation being triggered and how group members responded. This informal analysis was confirmed by a formal analysis of the pilot once the study was complete.

The sequence of adaptations triggered twice during the pilot: at 8 minutes into the session and again at 23 minutes. In both cases, the sequence progressed from *control lockout* to *voting* and was canceled before *escalate to authority*. This means that low-quality collaboration was still detected after at least one interval of *control lockout*, causing *voting* to be triggered. The group was able to sustain high-quality collaboration for at least three intervals while *voting* was active, so the adaptations were removed without messaging the teacher.

After reviewing the verbal and physical interactions among the students, we determined that the adaptations triggered appropriately and that they did appear to get the group to collaborate more effectively. However, the positive behavior change that occurred seemed to be a result of coercion—the adaptations proved so annoying that they forced the students to improve their working style without encouraging reflection. The main lesson learned was that it was too difficult to get the adaptations removed. Students did initially improve their collaboration but it degraded after a couple of minutes without feedback that they were on the right track. Additionally, the students read the warnings that appeared with each adaptation, but they complained that they felt they were already doing what was asked of them. When *voting* triggered a second time, they were quickly able to get back on track, maintain a shared focus but with little discussion or deep engagement with the content.

As a result of these observations from the pilot study, we reduced the length of time a group had to sustain high-quality collaboration in order to remove the adaptations from four minutes to two minutes. We also added an additional adaptation, *prompt* (Figure 1, top right), before *control lockout*. This adaptation simply provides students with a reminder of what is expected of them and gives them the opportunity to self-correct before the restrictive adaptations are triggered. The time that warning messages remained on-screen before activating *control lockout* and *voting* was reduced to 20 seconds as the video showed that to be enough time for the students to read the message. The rest of the implementation details remained the same.

Data analysis

Formal analysis of all study sessions began by coding the videos for social regulation using the same codes that were used to develop the collaboration quality detection approach (Evans et al., 2016). The bulk of the coding was carried out by a doctoral student who was unfamiliar with the adaptations being evaluated. To establish inter-rater reliability, the student and the first author independently coded a session from the baseline data collection phase of the study. Codes were applied to episodes (Chi, 1997), and each episode could contain multiple codes. The majority of codes had a Cohen's kappa (κ) above 0.61, typically considered "substantial" agreement, with several codes above 0.81, or "almost perfect" agreement (Landis & Koch, 1977).

To determine if the adaptations were triggered appropriately, we looked at the video codes in the intervals leading up to the triggering of each adaptation. If the video codes that occurred in that same interval were primarily negative, the adaptations were considered appropriate. We also looked for video intervals that showed improvement while adaptations were active to determine if the improvements were detected. Finally, we looked for video intervals that showed collaboration problems that were undetected by the computer.

To understand how students responded to the adaptations, we first looked at the video codes in intervals immediately following the activation of an adaptation. For an adaptation to be considered successful, the video codes should show improved collaboration. Due to the frustration observed in the pilot, we also reviewed the videos from the evaluation sessions to understand students' emotional responses to the adaptations.

Results

The results from the summative evaluation sessions suggest that the revised approach to triggering adaptations was more effective. Collaboration improved, along some dimensions, immediately following the first interventionist adaptation in the sequence (*prompt*) every time it was triggered, meaning that later adaptations were never triggered. This is a positive result because the *prompt* appeared to lead to more effective collaboration very quickly and consistently, but a side effect is that most of the adaptations were not therefore tested in this phase of the study. Additionally, the adaptations could not address all collaboration problems—disengaged students who showed no inclination to participate in the group work remained disengaged whether or not adaptations were present, and off-task interactions taking place away from the tabletop computer could not be detected. The biggest positive impacts of the adaptations were reduced disruption caused by individual students and less time spent engaged in low-quality collaboration.

Each group used the tabletop computer twice during the final evaluation phase of the study. Of the two groups that used the tabletop with adaptations available (Groups 1 and 3), only Group 1 triggered the adaptations. In Group 1's first session in the evaluation phase, they triggered the *prompt* once, at 4 minutes into the session. The video showed that the prompt was triggered at an appropriate time, after several minutes of the group being off task and pressing buttons on-screen without any explicit coordination. The prompt included a reminder that students should make sure they understood the task goal and that they were working toward it. When the prompt appeared, the group did revisit the instructions. The sequence of adaptations did not progress to the next stage because, immediately after reading the instructions, the group began to engage in on-task work. However, only some students in the group were engaged during most subsequent episodes. The disengaged, off-task students typically refrained from touching the computer after the prompt appeared in this session and were therefore undetectable. This behavior was noticeably different from Group 1's baseline and pilot sessions, in which these students would attempt to interact with the screen without fully engaging in the activity, disrupting students who were engaged.

In Group 1's second session in the evaluation phase, they triggered the *prompt* twice, at 9 and 17 minutes. The sequence did not progress beyond the *prompt* in either case. The video analysis showed that, in the intervals leading up to the first prompt, the whole group was engaging in primarily low-quality collaboration for around a minute, followed by a period of high-quality collaboration between two students with the other two students completely off task. After the prompt was dismissed, the off-task students continued to be off-task but refrained from touching the screen while the other two students worked collaboratively. In the intervals leading to the second instance of the *prompt*, one of the off-task students became interested in the screen, trying to take control of a particular object by repeatedly hammering on it. This interaction was highly disruptive to the engaged students, who were close to completing the task. When the prompt appeared for the second time, the engaged students appeared annoyed by it but they were able to dismiss it quickly and it deterred the other student from hammering on the screen. He sat back from the computer but made some verbal contributions to the collaboration—encouraging his teammates as they solved the assigned task.

We also compared the length of Group 1's periods of sustained low-quality collaboration taking place at the computer in three types of intervals: (1) "pre-adaptation"—intervals that cause an adaptation to trigger; (2) "adaptation active"—intervals in which an adaptation is active; and (3) "no adaptation"—intervals during which no adaptations are present. "Periods of sustained low-quality collaboration" means periods of time with one or more continuous episodes of low-quality collaboration.

Table 1 shows that the adaptations appeared to reduce the length of periods of low-quality collaboration involving the computer. The median length of periods of low-quality collaboration was consistently longer during intervals that caused an adaptation to trigger than during other intervals in all sessions where adaptations were present. In the baseline session, when the adaptations were not in use, the median length of sustained low-quality collaboration was at least twice that of the sessions where adaptations were available. In both final evaluation sessions, there were considerably more occurrences of low-quality collaboration when no adaptations were present than in other intervals. This effect occurred because, when there were no adaptations present, occurrences of low-quality collaboration were brief and punctuated by high-quality collaboration, causing the number of occurrences to increase and the length of the occurrences to decrease.

Table 1: The median length (in seconds) of sustained periods of low-quality collaboration in Group 1's sessions

Interval Type	**Baseline**		**Pilot**		**Evaluation 1**		**Evaluation 2**	
	Median length (s)	# of occurrences	Median length (s)	# of occurrences	Median length (s)	# of occurrences	Median length (s)	# of occurrences
Pre-adaptation			16	7	106	1	15	5
Adaptation active			9.5	19	0	0	11	2
No adaptation	20	11	9	6	8	19	10	15

Although Group 3 did not trigger the adaptations, the video analysis showed a serious collaboration problem—only one student engaged with the task while the other two group members sat back from the tabletop computer, chatting and using their phones. This highlights a known limitation of the detection approach used in our system, namely that it is only able to detect interactions with the screen (Evans et al., 2016).

Discussion and conclusion

The results show that our system was able to detect certain collaboration problems—primarily disruption caused by individual students and poor coordination among group members. Adaptations were triggered appropriately in these instances. How students in Group 1 received the adaptations differed by how motivated they were to engage with the activity—motivated students adjusted their behavior in a positive direction but already disengaged students were put off completely. Although this effect led to positive outcomes for the motivated students, who were able to make progress where they had previously been blocked by disruptive students, it was problematic for the disengaged students.

Although Group 3 never triggered the sequence of adaptations, both Group 1 and 3 saw the *group awareness icon* change color in response to the collaboration quality detected by the computer. However, with the exception of a single utterance in the pilot, there was no evidence that either group made use of the icon.

The *prompt* appeared to be effective at encouraging Group 1 to think about how they were interacting and to make sure they were working on the task as assigned. The first time the prompt appeared, the students took time to read it and follow its advice. The second and third time it appeared, the students were quicker to dismiss it, possibly due to familiarity, but both times, it caused an off-task student to stop disruptive behavior.

We consider the fact that, once *prompt* was added to the sequence of adaptations, no further adaptations were triggered, to be a positive outcome for this work. In all cases, the prompt was followed by sustained periods of high-quality collaboration, albeit only for those students who were engaged in the task. Beyond this study, the ideal outcome of using these adaptations over a longer period of time would be that they render themselves unnecessary—with such an outcome it would be possible to conclude that the adaptations successfully scaffold effective collaboration, fading once a group has adopted the principles that the adaptations support (Pea, 2004).

In this paper, we have described the design, implementation, and evaluation of a set of tabletop software adaptations to encourage effective collaboration when problems are detected. Presented as a sequence that grows increasingly restrictive if collaboration does not improve, our adaptations showed promise as supports for groups that struggle with disruptive behavior. Due to the small scale of our classroom field evaluation, further study is needed to determine the extent of our approach's effectiveness, but overall, our work demonstrates that tabletop applications that can detect and adapt to poor-quality collaboration can encourage more effective group work by deterring disruptive behavior.

References

Chi, M. T. H. (1997). Quantifying Qualitative Analyses of Verbal Data: A Practical Guide. *Journal of the Learning Sciences*, *6*(3), 271–315.

Dillenbourg, P., & Jermann, P. (2007). Designing Integrative Scripts. In F. Fischer, I. Kollar, H. Mandl, & J. M. Haake (Eds.), *Scripting Computer-Supported Collaborative Learning: Cognitive, Computational and Educational Perspectives*. New York: Springer.

Evans, A. C., Davis, K., Fogarty, J., & Wobbrock, J. O. (2017). Group Touch : Distinguishing Tabletop Users in Group Settings via Statistical Modeling of Touch Pairs. In *Proceedings of CHI 2017,* 35–47. New

York: ACM Press.
Evans, A. C., Wobbrock, J. O., & Davis, K. (2016). Modelling collaboration patterns on an interactive tabletop in a classroom setting. In *Proceedings of CSCW 2016,* 860–871. New York: ACM Press.
Evans, A., & Wobbrock, J. O. (2014). Filling in the gaps: capturing social regulation in an interactive tabletop learning environment. In *Proceedings of ICLS,* 1157–1161. Boulder, Colorado: ISLS.
Facing the Future, & Snow Leopard Trust. (2009). *Engaging Students in Conservation: Protecting the Endangered Snow Leopard*. Seattle, WA. Retrieved from http://www.snowleopard.org/learn/resources-for-educators
Hadwin, A. F., Oshige, M., Gress, C. L. Z., & Winne, P. H. (2010). Innovative ways for using gStudy to orchestrate and research social aspects of self-regulated learning. *Computers in Human Behavior*, *26*(5), 794–805.
Järvelä, S., & Hadwin, A. F. (2013). New Frontiers: Regulating Learning in CSCL. *Educational Psychologist*, *48*(1), 25–39.
Järvelä, S., Kirschner, P. A., Panadero, E., Malmberg, J., Phielix, C., Jaspers, J., Koivuniemi, M, & Järvenoja, H. (2014). Enhancing socially shared regulation in collaborative learning groups: designing for CSCL regulation tools. *Educational Technology Research and Development*, *63*(1), 125–142.
Kumar, R., & Rosé, C. P. (2011). Architecture for Building Conversational Agents that Support Collaborative Learning. *IEEE Transactions on Learning Technologies*, *4*(1), 21–34.
Landis, J., & Koch, G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*(1).
Malmberg, J., Järvelä, S., Järvenoja, H., & Panadero, E. (2015). Computers in Human Behavior Promoting socially shared regulation of learning in CSCL : Progress of socially shared regulation among high- and low-performing groups. *Computers in Human Behavior*, *52*, 562–572.
Martinez-maldonado, R., Kay, J., Yacef, K., Edbauer, M. T., & Dimitriadis, Y. (2013). MTClassroom and MTDashboard: supporting analysis of teacher attention in an orchestrated multi-tabletop classroom. In *Proceedings of CSCL,* 320–327. Madison, WI: ISLS.
Martinez-Maldonado, R., Yacef, K., Kay, J., Al-Qaraghuli, A., & Kharrufa, A. (2011). Analysing frequent sequential patterns of collaborative learning activity around an interactive tabletop. In *Proceedings of EDM,* 111–120. International Educational Data Mining Society.
Martinez, R., Collins, A., Kay, J., & Yacef, K. (2011). Who did what? Who said that?: Collaid: an environment for capturing traces of collaborative learning at the tabletop. In *Proceedings of ITS,* 172–181. New York: ACM Press.
Panadero, E., & Järvelä, S. (2015). Socially Shared Regulation of Learning: A Review. *European Psychologist*, *20*(3), 190–203.
Pea, R. D. (2004). The social and technological dimensions of scaffolding and related theoretical concepts for learning, education, and human activity. *Journal of the Learning Sciences*, *13*(3), 423–451.
Rogat, T. K. & Linnenbrink-Garcia, L. (2011). Socially Shared Regulation in Collaborative Groups: An Analysis of the Interplay between Quality of Social Regulation and Group Processes. *Cognition and Instruction*, *29*(4), 375–415.
Trausan-Matu, S., Dascalu, M., & Rebedea, T. (2014). PolyCAFe-automatic support for the polyphonic analysis of CSCL chats. *International Journal of Computer-Supported Collaborative Learning*, *9*(2), 127–156.
Volet, S., Vauras, M., & Salonen, P. (2009). Self- and social regulation in learning contexts: An integrative perspective. *Educational Psychologist*, *44*(4), 215–226.
Wang, H. C., Rosé, C. P., & Chang, C. Y. (2011). Agent-based dynamic support for learning from collaborative brainstorming in scientific inquiry. *International Journal of Computer-Supported Collaborative Learning*, *6*(3), 371–395.
Wise, A. F. (2014). Designing pedagogical interventions to support student use of learning analytics. In *Proceedings of LAK*, 203–211.

Acknowledgments
We thank Cindy Hmelo-Silver for her guidance on this project. This work was supported in part by the National Science Foundation under awards IIS-0952786 and IIS 1053868. Any opinions, findings, conclusions or recommendations expressed in this work are those of the authors and do not necessarily reflect those of the National Science Foundation.