

# Filling in the Gaps: Capturing Social Regulation in an Interactive Tabletop Learning Environment

Abigail Evans and Jacob O. Wobbrock, Information School, University of Washington, Seattle, USA  
Email: abievens@uw.edu, wobbrock@uw.edu

**Abstract:** A study of small groups collaborating at an interactive tabletop was conducted. Group discussions were coded according to the type and quality of social regulation processes used. Episodes of high and low quality social regulation were then matched with the software logs to identify patterns of interaction associated with quality of social regulation. A key finding is that instances of low-quality social regulation were characterized by more than twice as much interaction with the software as high-quality instances.

## Introduction

Creating software that can adapt to the needs of learners and create personalized learning environments is an oft-cited goal of researchers and designers of educational technology (Jermann, Mühlenbrock, & Soller, 2002; Lonchamp, 2010; Marcos-García, Martínez-Monés, Dimitriadis, & Anguita-Martínez, 2007; Martínez, Collins, Kay, & Yacef, 2011a). Collaborative learning environments represent a particular challenge in personalization, as they must take into account interactions between multiple learners. Collaborative learning at tabletop computers brings further challenges, as much of the interaction between learners takes place face-to-face, away from the shared interface, and differentiating between individuals can be difficult (Dillenbourg & Evans, 2011).

The process of small group face-to-face collaborative learning is complex and many factors contribute to learning outcomes. The effectiveness of the group learning experience can be influenced by the personalities and motivations of individual students, their background knowledge, the nature of the task, and even their time management skills. Students do not always know how to collaborate in a manner that is productive and conducive to learning for all members (Rogat & Linnenbrink-Garcia, 2011). In the classroom, the teacher will often circulate among groups, but she can only visit one group at a time, meaning that struggling groups may not receive the help they need in a timely manner. Software that can adapt to a group's collaboration needs and abilities in real-time could prove useful in filling in the gaps while the teacher spends time with other groups.

The goal of this research is to uncover what physical interactions with a tabletop computer reveal about the collaborative interactions taking place above the table, which is arguably where the learning actually occurs. Rather than trying to determine if a group is meeting content-based learning objectives, our intention is to understand the collaboration itself to inform future development of tabletop software that can adapt to scaffold effective collaboration. While it may not always be true in practice (Dillenbourg & Evans, 2011; Rogat & Linnenbrink-Garcia, 2011), an implicit assumption in this work is that effective collaboration leads to positive learning outcomes, while ineffective collaboration hinders learning.

Important to this work is 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). Rogat and Linnenbrink-Garcia (2011) showed social regulation to be very important to small group collaboration. They identified 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. We use this conception of social regulation to guide the analysis of our participants' collaboration.

## Related Work

When learners collaborate at an interactive tabletop computer, only interactions with the software are usually captured. The verbal and gestural interactions among students that represent the learning process are not visible to the computer, making it tricky to capture meaningful assessment data or to provide relevant dynamically generated feedback to learners and teachers. A small body of work has begun to explore this problem space.

Martínez *et al.* (2011a; 2011b) created visualizations of learners' verbal contributions and physical interactions with the software during a group concept-mapping activity at a tabletop computer. The visualizations proved useful for determining how equitable the group members were in terms of the quantity of participation, but could not communicate the quality of the collaboration itself. The same authors (2013) later created a dashboard that helped teachers determine which groups needed attention as they worked by enabling the teacher to see how closely each group's concept map matched that of the teacher. Martínez *et al.*'s positive results suggest that this is an effective strategy for activities in which students work towards a particular ideal, or well-defined solution. It would not, however, be effective in the absence of an expert model, as is generally the case in creative work, or in collaborative work in the spirit of knowledge-building (Scardamalia & Bereiter, 1994), in which learners build new knowledge and understanding. Our work seeks to model and adapt to collaborative processes in these more learner-led contexts.

Fleck *et al.* (2009) explored the relationship between students' discussions and their actions on an interactive tabletop. They found that verbal elements of successful collaboration, such as making and accepting suggestions or negotiating, were often complemented by particular actions in the software. This finding suggests that some aspects of collaborative discussion should be detectable through interaction analysis. Our work builds on these findings to identify patterns of action over time that can reveal groups' social regulation processes.

## The Study

Eleven adults (7 female, 4 male) worked in small groups (two dyads, a group of three, and a group of four) to analyze and compare two poems, *Birches* by Robert Frost (1969), and *Fern Hill* by Dylan Thomas (2003). We chose poetry analysis for our study because many students find it challenging, and yet it requires little background knowledge, making it an authentic learning activity for most participants. Poetry analysis is interpretative, which also makes it an ideal activity for small group work as students can reach deeper levels of understanding by externalizing their own thoughts and building on the ideas of others. The participants were given 30 minutes to work on the following task adapted from McMahan, Funk, & Day (1998):

### Compare and contrast two poems, *Fern Hill* by Dylan Thomas and *Birches* by Robert Frost.

Answer the following questions and support your answers with evidence from the text. As you answer each question, consider how the two poems are similar and how they are different.

1. What is the **theme** (the central idea) of each poem?
2. Who is the **speaker** in each poem? How would you describe the speaker?
3. What **imagery** does each poet use? How do the images relate to each other and contribute to the theme?



Figure 1. A screenshot of the software used in the study

The participants used a Microsoft PixelSense, which has a 40" multitouch screen and can comfortably seat four adults around it. The software (see Fig. 1), created specifically for this study, was intended to be used as a tool to support discussion rather than for creating a result to be turned in. At the end of the activity, the participants were asked to verbally summarize their findings to a researcher as if they were feeding back to the rest of a class after a period of small-group work. This meant that the participants were relatively unconstrained in how they tackled the task and were free to make use of the software in the ways they found most useful.

The text of both poems was presented on-screen and supplemented by audio recordings of readings by the poets. Participants could take notes by clicking buttons to add "notecards" to the screen. Notecards were color-coded according to each element of the poems they were asked to discuss (*e.g.*, red for theme, green for speaker, and yellow for imagery) and participants could add as many notecards as they wanted. Each notecard contained a text field in which participants could type an observation about the text, and an area for collecting supporting evidence by dragging and dropping lines from the poems onto the notecard. Once dragged onto a notecard, individual words could also be highlighted. Participants could annotate the poems directly by drawing on them with their fingers. The software also included a set of optional prompt questions to stimulate discussion. Each tool described so far could be freely moved around the screen. In addition, fixed pieces of "scratch paper" were provided at each corner to enable each participant to take their own notes.

Two video cameras recorded the discussions. Interactions with the tabletop were recorded in log files detailing when and where a touch happened, and what action was performed. The researcher left the room while the participants worked to encourage them to give an authentic summary at the end of the activity.

## Data Analysis

To determine how social regulation manifested in each group, each video was coded along the three dimensions of social regulation described by Rogat and Linnenbrink-Garcia (2011), and for high, moderate, and low quality.

The codes are listed in Table 1, but detailed descriptions of each code, as well as high and low quality examples, can be found in the authors' article.

Table 1: Descriptions of social regulation codes, excerpted from Rogat and Linnenbrink-Garcia (2011)

<i>Codes</i>	<i>Description</i>
Social regulation	Group efforts to regulate their conceptual understanding, task work, and engagement.
Planning	Reading and interpreting task directions, designating task assignments, discussing how to go about solving the problems.
Monitoring	Evaluating content understanding, the shared product, assessing progress, or plan for completing the task.
Behavioral engagement	Encouraging an off-task group member to re-engage, reminding a group member to return to task.

In order to draw useful insights from the log files of touch data, individual touches needed to be grouped into sequences that represented purposeful actions—actions that served a particular purpose. For example, imagine a participant wants to read one of the poems, which is currently located closer to another participant on the other side of the screen. He makes one touch to move the poem closer to himself, followed by a second touch to rotate it so it is oriented properly from his perspective. In this case, the two touches together represent a single purposeful action. Three features are needed to make that determination: (1) who carried out the touch (the owner); (2) the object that was touched; and (3) the timing of the touches. Items 2 and 3 are readily available in the logs but, as the PixelSense cannot natively differentiate individual users, for this study, the owner was manually labeled by matching the timestamp of the touch to a frame from the video.

For the first group, touches were grouped into meaningful actions by close inspection of the log file alongside the video. The timing between touches was studied to determine the distinction between sequences of touches that represented purposeful actions, and sequences that should be considered separate actions. Clear rules emerged from this analysis for sequences of touches owned by a single person in terms of object type and length of time between touches (see Table 2).

Table 2: Rules for grouping touch event data into purposeful actions by time between touches

<i>Objects touched are:</i>	<b>Purposeful action sequence</b>	<b>Separate actions</b>
<i>Related</i>	$\leq 50$ seconds	$> 50$ seconds
<i>Unrelated</i>	$\leq 15$ seconds	$> 15$ seconds

The object types *related* and *unrelated* refer to objects' function in the context of the activity. For example, the time between a touch on the Thomas poem followed by a touch on a line from the Thomas poem that has been dragged to a notecard would be categorized as *related*, while a touch on the Frost poem followed by a touch on the instructions would be *unrelated*. Objects could be related either by poem, *e.g.*, a notecard containing lines from Thomas is related to the Thomas poem; or by task sub-question, *e.g.*, a theme notecard containing lines from Frost is related to the Thomas poem in "Annotate Themes" mode.

For groups 2, 3, and 4, logged touches were automatically grouped according to the rules given in Table 2. The rules were then validated by repeating the video inspection for these groups and, for each category, calculating the percentage of touches that were correctly categorized. The rules proved to be highly accurate, ranging from 87.2% to 100.0% correct.

The screen of the PixelSense is highly sensitive and as a result, the touch logs contained a lot of noise, often due to participants catching on-screen objects with their elbows while reaching for another object. Another source of noise came from multi-touch input—using more than one finger to carry out a single action, such as rotating an object using multiple fingers. Simple heuristics were used to filter out these types of noise: (1) touches overlapping (in time) by the same person on the same object were combined, (2) when there were overlapping touches by the same person on different objects and one touch was close to the edge at which the person was sitting, the edge touch was removed, and (3) when there were overlapping touches on multiple objects located directly on top of each other, the top touch was kept and all others were removed.

The third stage of analysis brought together verbal and physical interactions. The goal of this stage was to find patterns of touches that reflected social regulation processes used by each group, as well as their quality. Therefore, we identified the processes for which there were clear differences among the groups in terms of quality. This substantially narrowed the focus of the subsequent analysis as the groups had more similarities than differences. Timestamps from the video that delineated episodes where a group was engaging in a particular regulatory process were applied to the pre-processed touch data. The episodes were then compared to see whether patterns emerged. Only touches to the shared objects (*i.e.*, everything but individuals' separate "scratch paper") were considered at this stage of the analysis.

## Results: Forms and Quality of Social Regulation

All groups spent around ten minutes of the activity silently reading the poems and listening to the recordings before starting discussion. The nature of the task demanded relatively little *planning* and all groups were coded as high quality in this area.

After the initial period of preparation, however, each group took a different approach in their discussions and their use of the software. Group 1 (3 participants) structured their discussion around the task's three sub-questions, comparing the two poems as they considered the themes, speakers, and imagery in order. They took notes using the notecards and copied lines from each poem to support their observations. Group 1 made the most effective use of the allotted time, covering each aspect of the task in depth. Their *monitoring* processes were consistently high quality across all 3 sub-codes.

Group 2 (4 participants) did a great deal of comparing and contrasting of details in the poems. They were the only group to make use of the optional discussion prompts. They annotated the poems directly and wrote notes on their individual virtual notepaper. Group 2 did not follow a clear structure. They covered each sub-question, but not in a linear fashion. They exhibited consistently high-quality *monitoring* of the plan and progress. Although their content monitoring was mostly high quality, there were some low-quality instances, largely due to one group member's occasionally dismissive and unresponsive treatment of others' contributions.

After reading both poems, Group 3 (a dyad) decided to tackle one poem at a time. They started with the Thomas poem and covered each task sub-question in turn with just that poem, taking notes using the notecard feature for the theme and speaker before switching to directly annotating the imagery in the poem. They did not get to the Frost poem until near the end of the activity, eventually running out of time having only considered its theme and not its speaker or imagery. As a result, the group was not able to make many comparisons between the two poems, although their analysis of *Fern Hill* was exceptionally thorough. In terms of social regulation, their discussion was focused on content monitoring, all of which was high quality. In contrast, progress monitoring was largely missing and monitoring of the plan was very limited, though generally high quality.

Group 4 (a dyad) were more focused on organizing the virtual workspace and collecting lines than on actually engaging with the poems. Their discussion of both poems, though valid, was mostly superficial and more of a summary than an analysis. Most instances of content monitoring were of low quality, although they did begin to improve near the end of the discussion. They were easily distracted by superficial concerns, which meant that thoughtful contributions were often unheeded—an indicator of low-quality progress monitoring. They did, however, also engage in high-quality progress monitoring, regularly tracking what remained to be done. Their plan monitoring was generally moderate as, although they frequently sought to clarify the instructions and evaluate their plan, their continual monitoring of the plan hindered their ability to enact it.

## Results: Social Regulation in the Touch Data

As described above, the groups differed in quality along the monitoring dimension of social regulation, so the three monitoring sub-codes were used to identify patterns of touch data that reflected quality of monitoring processes. Results were normalized by group size to facilitate comparison.

### Monitoring Content

Monitoring content was the most-used process of social regulation, with noticeably longer duration than the other processes. Quality variations in content monitoring revealed distinctive patterns of touches. Low-quality episodes, all from group 4, were characterized by sequences of purposeful touches to unrelated objects. Group 4, which was consistently low quality, averaged 23.5 such sequences, more than double that of the high-quality groups. Group 1 averaged just 2.3, while group 2 averaged 10.5, and group 3 averaged 11.0. During low quality episodes, sequences of touches were also much temporally denser—many touches in quick succession. Overall, group 4 spent more than twice as much time, 476 seconds per person, touching the screen than any other group. The high-quality groups were similar to each other: group 1 = 199s, group 2 = 136s, and group 3 = 190s. Finally, high-quality episodes were characterized by only one person interacting with the screen at a time, moving from one person to another with little overlap, suggesting turn-taking. In groups 1, 2, and 3, overlapping touches by two or more people made up 10.5%, 10.7%, and 9.6% of their total touches to shared objects, respectively. Conversely, low-quality episodes often featured overlapping touches. For group 4, overlapping touches made up 23.6% of their total touches to shared objects.

### Monitoring the Plan

The log data contained indicators that plan monitoring was occurring. Participants sometimes touched the instructions on-screen when revisiting the task and re-evaluating their plan and otherwise did not interact with it. However, the variations in quality among the groups did not present distinctive patterns.

### Monitoring Progress

Episodes of progress monitoring did not show any distinctive patterns in the touch data, regardless of quality. Progress monitoring episodes were generally brief, predominantly verbal interactions that did not generate much

touch data. While deictic gestures did appear in the video, they were generally above the screen and so were not captured by the log data.

## Conclusion

High-quality content monitoring has been shown to be important in the construction of shared meaning; therefore, the ability to automatically detect the quality of content monitoring in a collaborative tabletop environment would be extremely useful. While the clear distinction between the high- and low-quality episodes is a promising result, it is important to note that the low-quality episodes were limited to just one group and therefore could be idiosyncratic to their particular style of interaction. However, the similarities between high-quality episodes across groups suggest that the patterns found might be generalizable.

The prevalence of episodes of content monitoring significantly aided the process of finding obvious patterns. In the context of the poetry analysis activity, there was relatively little need for planning or monitoring planning as the task structure was simple, so the episodes were brief and few. Also, as the task did not require the production of a shared artefact to be submitted for evaluation, there was perhaps less need for progress monitoring. It would be interesting to see if activities with different structures evoke different social regulation behaviors, which might cause more patterns to emerge. Additionally, the adult volunteers participating in this study remained on-task and engaged, so regulatory processes related to behavioral engagement were not employed. We expect this would not be the case in a real classroom environment. We are planning to conduct a follow-up study with middle and high school aged children to investigate these questions, and also to see whether the content monitoring patterns discovered transfer to other domains and activities, along with the rules we used to identify purposeful touch sequences.

## References

- Dillenbourg, P., & Evans, M. (2011). Interactive tabletops in education. *International Journal of Computer-Supported Collaborative Learning*, 6(4), 491-514.
- Fleck, R., Rogers, Y., Yuill, N., Marshall, P., Carr, A., Rick, J., & Bonnett, V. (2009). Actions speak loudly with words: Unpacking collaboration around the table. *Proceedings of the International Conference on Interactive Tabletops and Surfaces* (189-196). New York: ACM Press.
- Frost, R. (1969). *Birches. The Poetry of Robert Frost*. New York: Henry Holt and Co.
- Jermann, P., Muhlenbrock, M., & Soller, A. (2002). Designing computational models of collaborative learning interaction. *Proceedings of the Conference on Computer-Supported Collaborative Learning* (730-732). Boulder: International Society of the Learning Sciences.
- Lonchamp, J. (2010). Customizable computer-based interaction analysis for coaching and self-regulation in synchronous CSCL systems. *Journal of Educational Technology & Society*, 13(2), 193-205.
- Marcos-Garcia, J. A., Martinez-Mones, A., Dimitriadis, Y., & Anguita-Martínez, R. (2007). A role-based approach for the support of collaborative learning activities. *e-Service Journal*, 6(1), 40-58.
- 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. *Proceedings of the Conference on Computer-supported Collaborative Learning* (320-327). Madison: International Society of the Learning Sciences.
- Martinez, R., Collins, A., Kay, J., & Yacef, K. (2011a). Who did what? Who said that? Collaid: An environment for capturing traces of collaborative learning at the tabletop. *Proceedings of the International Conference on Interactive Tabletops and Surfaces* (172-181). New York: ACM Press.
- Martinez, R., Kay, J., & Yacef, K. (2011b). Visualisations for longitudinal participation, contribution and progress of a collaborative task at the tabletop: *Proceedings of the Conference on Computer-Supported Collaborative Learning* (25-32). Sydney: International Society of the Learning Sciences.
- McMahan, E., Funk, R., & Day, S. (1998). *The elements of writing about literature and film*. London: Longman.
- 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.
- Scardamalia, M. & Bereiter, C. (1994). Computer support for knowledge-building communities. *Journal of the Learning Sciences*, 3(3), 265-283.
- Thomas, D. (2003). Fern Hill. *The Poems of Dylan Thomas*. New York: New Directions.

## Acknowledgements

This work was supported in part by the National Science Foundation under grant IIS-0952786. 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.