

# Challenging Assumptions: using sliding window visualizations to reveal time-based irregularities in CSCL processes

Gregory Dyke, Carnegie Mellon University, Pittsburgh, PA, USA, gregdyke@gmail.com  
Rohit Kumar, Raytheon BBN Technologies, Cambridge, MA, USA, rohit.kumar@gmail.com  
Hua Ai, Georgia Tech, Atlanta, GA, USA, hua.ai@cc.gatech.edu  
Carolyn P. Rosé, Carnegie Mellon University, Pittsburgh, PA, USA, cprose@cs.cmu.edu

**Abstract:** Temporality and the unfolding of collaborative processes over time are an important aspect in understanding the underlying mechanisms of collaborative learning. In this paper, we describe how interactive sliding window visualizations can integrate with existing methodologies to provide new insight into how interactions develop over time. We illustrate the value of these visualizations by performing a secondary analysis of a corpus analysed in a previous study to show how new insight can be gained and how results at a scale of a whole interactive session can be challenged when looking for time-based irregularities. We conclude by discussing the kinds of insights that can be gained from sliding window visualizations, how they overcome certain methodological limitations of existing techniques for dealing with temporality in CSCL and, in turn, the methodological limitations which they introduce.

## Introduction

In recent years, many authors and collaborations have addressed the problem of temporality in Computer Supported Collaboration (Reimann, 2009; Kapur, 2011; Wise & Chiu, 2011; Suthers, Teplovs, De Laat, Oshima, Zeini, 2011). The variety of beliefs (Suthers, 2006) as to the mechanisms that render collaboration beneficial for learning require that these collaborative processes be studied in detail. The irregularities of these processes are of fundamental importance in CSCL for two reasons. First, learning is primarily about change: acquiring or creating new knowledge and skills. Second, the introduction of technological artifacts to mediate collaboration allows us to study the ways in which collaborative meaning-making is affected by the environment, in an appropriation process which is expected to change over time.

Chiu & Khoo (2005), Jeong (2008), and Reimann, Frerejean, & Thompson (2009) have described and successfully employed a variety of statistical techniques which apply to sequences of events and allow the identification and characterization of recurrent patterns in interaction. While these techniques have revealed fascinating results, allowing time to be taken into account rather than factoring it out as an intractable element, they have some weaknesses. They are essentially concerned with identifying regularities at all points in time and must be creatively adapted to show change over time. Furthermore, the timescale of these regularities (if and when they exist) is at that of the event, or of the adjacency pair, and such granularities fail to account for the prior history of the collaborative and learning processes. Last, many social processes related to learning and collaboration, are not expected to happen at the smaller timescales (McGrath and Tschan, 2004), but occur concurrently at multiple timescales.

Borrowing techniques from time-series analysis and as suggested by several authors in CSCL (e.g. Muukkonen et al., 2007; Reimann, 2009) interactive time-series visualizations, calculated from sliding window analyses can make a methodological contribution to understanding the irregularities of the unfolding of collaborative processes over time. A sliding window analysis takes an indicator which can be calculated from a set of events, and rather than producing a summary value for a whole period of interaction, computes this indicator for a “window” of a given size (e.g. 5mn) which is then slid over the period of interaction by small chunks (e.g. 10s), producing the value of the indicator for each window, and showing how it changes over time.

In the remainder of this paper, we examine in greater detail the existing techniques for taking time into account and present a methodology based on interactive sliding window visualizations. We illustrate this methodology by performing a secondary analysis on a corpus analysed in a previous publication (Ai, Kumar, Nguyen, Nagasunder, & Rosé, 2010) and highlight the ways in which such visualizations can lead to a better understanding of the data, give a more subtle account of existing results, and allow new hypotheses and variables to be suggested and confirmed by triangulation using other methodologies. We conclude by discussing the theoretical and practical limitations of such visualizations, the affordances that make their interpretation easier, and outline their relevance for real-time automated analysis and support for computer-mediated collaboration.

## Identifying temporal (ir)regularity

The (stereo)typical coding and counting approach of Content Analysis (De Wever, Schellens, Valcke, & Van Keer, 2006) can be criticized for its adaptation from fields where temporality is less central, reducing interactions to summary indicators which no longer have a temporal component. Such indicators or summary

statistics, have nevertheless proven useful, particularly in hypothesis testing scenarios where a set of interactions are designed to vary only in the variables isolated for study. Indeed, time is not entirely lost as the act of manually coding assumes the situatedness of each turn or event being coded. However, removing the temporal aspect can lead to misattributions being made when statistical associations are found between counts over whole sessions and outcome measures, because the distribution in behaviors over time is not taken into account. In some of our recent work, for example, a certain form of scripting was found to be ineffective in producing a positive learning outcome (Dyke, Adamson, Howley, & Rosé, 2012). A closer examination of the temporality revealed that it produced the desired interactions (which might theoretically have produced a positive learning outcome), but these interactions disrupted the natural flow of interaction to such a degree that the benefits were balanced out.

Methods adapted from Conversational Analysis (e.g. Suthers, Dwyer, Medina, & Vatrapu, 2010) have had the opposite problem, with their costliness in analytic time making it hard to identify generalizations, and have led to the refinement of methods of describing collaborative processes in ways that focus as much on the relationships (or contingencies) between events as on the events themselves. Reimann (2009) argues that event-driven approaches that seek to model sequential process can provide a methodological link between the “thick” descriptions produced by fine-grained analyses and the robust generalizations of quantitative studies. He also provides a detailed overview of the kinds of techniques which could be used, along with their methodological implications.

Jeong (2005), Reimann et al. (2009), Thompson, Kennedy-Clark, Markauskaite, & Southavilay (2011) use Markov models to model collaborative processes and describe distinctive features which characterise certain conditions or certain populations of learners. Chiu & Khoo (2005) and Kapur (2011) describe statistical techniques which introduce lag-variables to examine the effect that the preceding turn (at a lag of -1) or older turns (at lags of -2, -3, etc) have on the current turn. While these approaches can be generalized to the non-linear structure of artefact-mediated interactions, they nevertheless tend to only apply to small time scales and to the search for regularities. In a study by Jadallah et al. (2011), lag variables ceased to have a discernable effect after seven turns. Even short-term effects, such as the introduction of technology for a one hour session would be expected to produce a gradual effect over the course of several hundred turns. As Reimann (2009) highlights, learning processes happen at multiple timescales, making it challenging to find new coherent temporal units over which larger-scale variables can be meaningfully computed.

Chiu & Khoo (2005) have used breakpoint analysis to identify focal points in interaction in which the nature of a variable shifts, marking, for example, a transition from a period of low to a period of high micro-creativity. Unfortunately, such statistical techniques are computationally expensive, both requiring datasets with many events, and limiting the number of breakpoints that the model can attempt to find. Mayfield, Rudnicky, & Rosé (2012) have used dynamic time warping (a technique for aligning temporal processes which are not of the same duration or do not happen at the same speed) to characterize differences between populations performing a given collaborative task. Jain, McDonough, Gweon, Raj, & Rose (2011) have used a dynamic bayesian network models to show how linguistic stylistic accommodation happens over the duration of an interaction.

These techniques, along with techniques developed for mining collaboration patterns in Educational Data Mining (e.g. Maldonado, Yacef, Kay, Kharrufa, & Al-Qaraghuli, 2011) tend to posit a determinism which may not be present or which may be influenced by too many variables which can neither be measured nor controlled. In our work, we take an Exploratory Sequential Data Analysis (Sanderson & Fisher, 1994) approach, providing the means to examine an overview of trends over time (and over multiple time periods) in order to quickly gain an appreciation for how a given interaction unfolds, what might be characteristic of a particular class of interactions and what might distinguish different classes of interactions (e.g. in different conditions, or with differing degrees of success).

## Sliding Window Visualizations

Sliding window visualizations extend the ideas of indicators computed as a summary of an interaction by reintroducing the notion of time. If we take a very simple indicator (e.g. distribution of turns among participants), this initially produces a summary of the distribution of turns over a whole interaction (e.g. telling us that all participants participated equally), but giving us no idea of change. By taking the smallest possible time-period (1 turn), the values of this distribution become very noisy (during a single turn, a single participant has 100% of the participation and the identity of this participant changes at each turn), telling us, that, at this micro level, interactions are very irregular. During a small portion of an interaction, however, two of the participants might be monopolizing the discussion, producing a 45/45% distribution for them, with the remaining 10% of turns contributed by the remaining participants. Sliding window visualizations show how the distribution changes over time throughout the interaction. As this window slides over times, it gradually “leaves” certain events out of the calculation of the indicator, and “introduces” new ones. At each iteration, the indicator is calculated for that window and plotted. The size of the window produces a certain amount of

smoothing or “blurring”, with one extreme being the window over the whole interaction (equivalent to the usual summary statistics) and the other being the jumpiness of the value per event.

Our approach differs from other visualization approaches (e.g. Jermann, Mullins, Nuessli, & Dillenbourg, 2011; Hmelo-Silver, Chernobilsky, & Nagarajan, 2009; Suthers et al., 2010) in that we automatically make the transition in granularity from the turns or events which are present in the raw data. In order to compensate for this (as there is no reason that there should be any comparability in terms of interactive and learning processes between two windows just because they contain the same number of events or the same amount of time), we provide the means to change the amount of smoothing (the size of the window) on the fly. Furthermore, we implement these visualizations within the Tatiana analysis framework (Dyke, Lund, & Girardot, 2009; Dyke, Lund, & Girardot, 2010), which provides synchronization between visualizations, allowing a human analyst to transition from the visualization to the raw data and back. In this way, any salient irregularities can be examined, verified to not be an artifact of smoothing at a particular granularity, and be used as a means of diving into the raw data. The interactive nature of our visualizations distinguish them from existing work in CSCL which has examined change over time (e.g. Wang & Rosé, 2007).

In accordance with the Tatiana data model, each event has a certain number of properties, including, when applicable, a *contents*, which is the message, utterance or text for that event. In order to introduce a certain amount of genericity in these visualizations, any ad-hoc indicator can be computed, with default indicators showing number of turns and number of words, as distributed, not only among participants (one of the key properties of an event), but also among other properties, such as a set of codes (e.g. ontask and offtask). We furthermore provide the means to view each visualization either in absolute numbers (A had 5 turns, B had 10) or in proportions (A had 1/3, B had 2/3).

## Social Conversational Agents for Scaffolding CSCL

To illustrate the ways in which such visualizations can be exploited, we will perform a secondary analysis on the dataset analyzed by Ai et al. (2010). This research was conducted to investigate how conversational agents (chatbots) can be used, not only to script the content of a collaborative discussion task, but also to provide social support to groups (such as showing solidarity, precipitating tension-release, and agreeing). It also investigates the effect of bias and alignment with student goals, in a situation where students are given competing goals. The study builds on an activity for which it had already been shown that providing social support produced a significant effect on learning (Kumar, Ai, Beuth, & Rosé, 2010). In this activity, students must design a Rankine cycle (a type of power plant), working in pairs to understand the trade-offs between power output and environmental friendliness.

The study was conducted in a 3x3 factorial design across 53 pairs of students in which two aspects of the conversational agents were manipulated. The first was the amount of social support: None, Low (at most 15% of agent turns are social), and High (at most 30%). The second was the goal-alignment: Green, Power, and Neutral. In all conditions, the agents guide the students through a carefully designed script going through a series of 4 KCDs (*Knowledge Construction Dialogues*), each of which discusses a parameter of the cycle and its implications: Tmax (maximum temperature), Pmin, Pmax (minimum, maximum pressure) and fuel. The wording of these dialogues is slightly different in the Green and Power conditions to highlight the influence of these parameters on environmental friendliness and power output respectively. For example, where the Green agent might say “What is bad about increasing the heat input to the cycle is that it increases the heat rejected to the environment.” The neutral tutor would simply say “Increasing heat input to the cycle increases the heat rejected to the environment.” After each KCD, the tutor prompts the students to discuss which value they would like to choose for that parameter and gives them a few minutes to conduct this discussion. At the end, the tutor concludes by prompting the students to review their choices for each parameter and come up with a final design. Students worked in pairs, with each pair having a student whose goal was to maximize power output and a student whose goal was to maximize environmental friendliness. Thus, each student can be considered to either be in the Match, Mismatch or Neutral condition, depending on their goal and that of the agent.

In the original analysis, the learning gains between pre- and post-tests were evaluated across each of the 9 conditions, as was the amount of offtask behaviour, the amount of agent abuse (the agent sometimes makes remarks at inappropriate times, leading the students to comment on how much they disliked it), and the amount of social behaviour. The main results were:

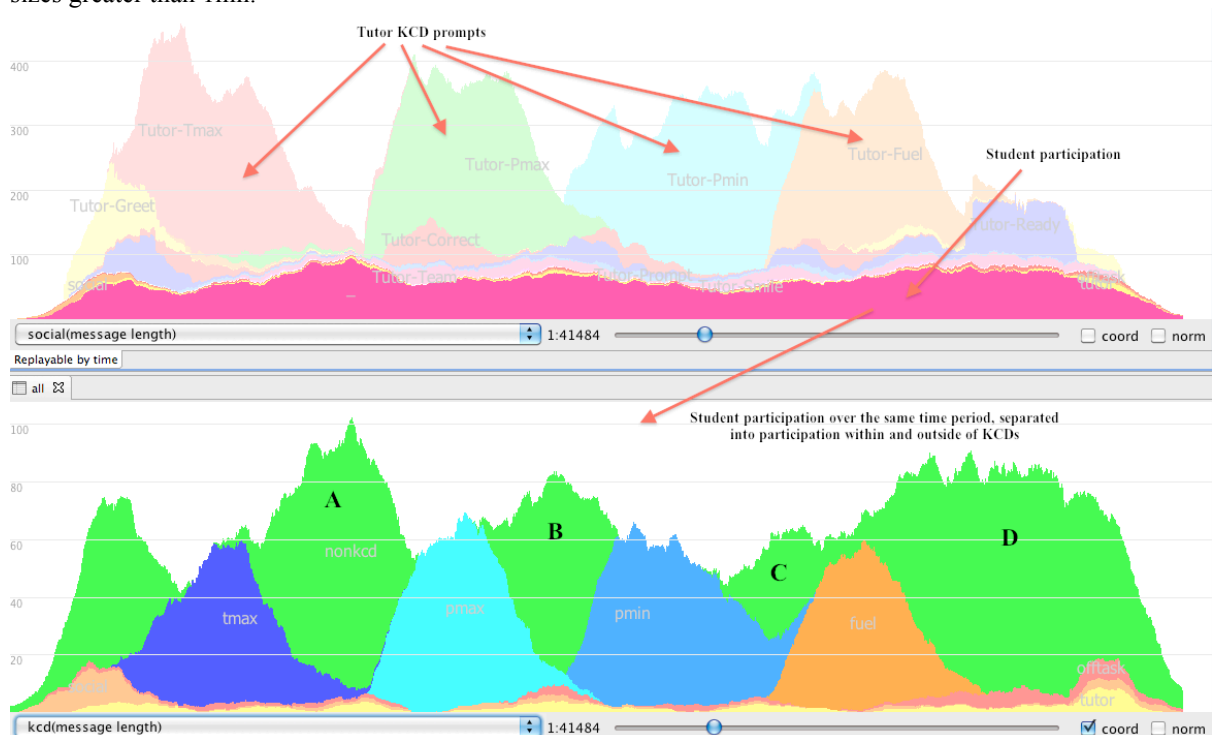
- Learning was significantly better in the low social condition than in the other two (particularly in the Match condition), with an effect size of .83 standard deviations.
- Social behaviour was significantly higher in both of the social conditions than in the None condition (effect sizes of 1.8 and 1.2)
- There was significantly more offtask behaviour in the None social condition than either of the other two (effect sizes of .35 in both cases)
- There was significantly more tutor abuse in the High social condition than in the Low or None condition (effect sizes of 1.1 and 1.28 respectively)

## A Secondary Analysis Looking for Irregularities

We now show how creating various forms of sliding window visualizations on this corpus allows us to reexamine the results presented above and to identify new insights and irregularities to produce new statistically significant results. More specifically, we would like to identify whether the low social condition affected participation over time in some way which can explain the increase in learning. Furthermore, we would like to examine the extent to which the social, offtask and tutor abuse behaviors were uniform throughout the conversation or whether they occurred at certain specific times.

### Single group – All groups

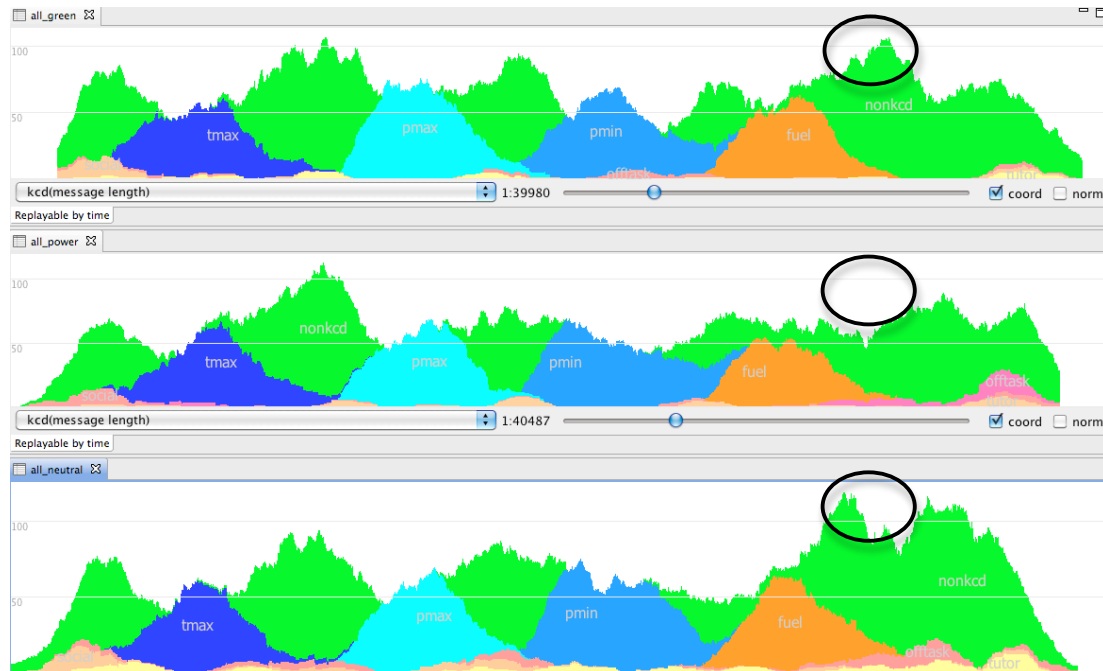
While the visualization of a single group gives a good idea of what happened in that group, it is very hard to get a feel for what is “normal” across 53 different groups. For this reason, and because the timing of the agent’s scripting was reasonably consistent (to within about 10s) across all groups, we combined all the sessions into a single time-aligned transcript and visualized it to get an idea for what is normal across all groups. Figure 1 (top) shows the general trends of how the conversations flowed over time (horizontal axis, each session being around 40mn in length), in words per minute (vertical axis), with tutor participation (pastels) separated according to the KCD or the type of social turn to which it belongs – showing clearly the flow from one KCD to the next, with minima in tutor participation in between. The student participation (pink, yellow and orange) shows noticeable peaks during the times when the tutor is not in a KCD. As a result of this observation, we coded each student utterance according to where it fit in the discussion (offtask, social, tutor abuse, on task in the Tmax, Pmax, Pmin, Fuel KCDs, or on task outside of one of the KCDs), producing Figure 1 (bottom), which hides all tutor turns and amplifies the scale of the student participation (the pink above is the sum of all the colors below), showing clearly 4 equal bouts of interaction with the tutor within KCD, and showing high participation in the interval after the first KCD (A), slightly lower participation after the second KCD (B) and lowest participation after the third KCD (C), before rising after the last KCD and during the final stages of the decision (D). These regularities (irregularities over time, regularities across groups), while not particularly surprising (students do a lot of reading during KCDs and participate less – there is also less to discuss as the power plants’ parameters become increasingly fixed), were not known to us previously. We also examined various student participation ratios: green vs. power; most learning vs. least learning; match vs. mismatch, but all these revealed near equal participation, throughout the interaction. Concerned that this might be a smoothing effect due to cumulating all groups over time, we visualized a more precise indicator which calculates within group ratios – and again obtained the same results. Although the match condition produced a slight effect on learning, it apparently produced no effect on participation over time: a consistent 50-50 split in student participation at all window sizes greater than 1mn.



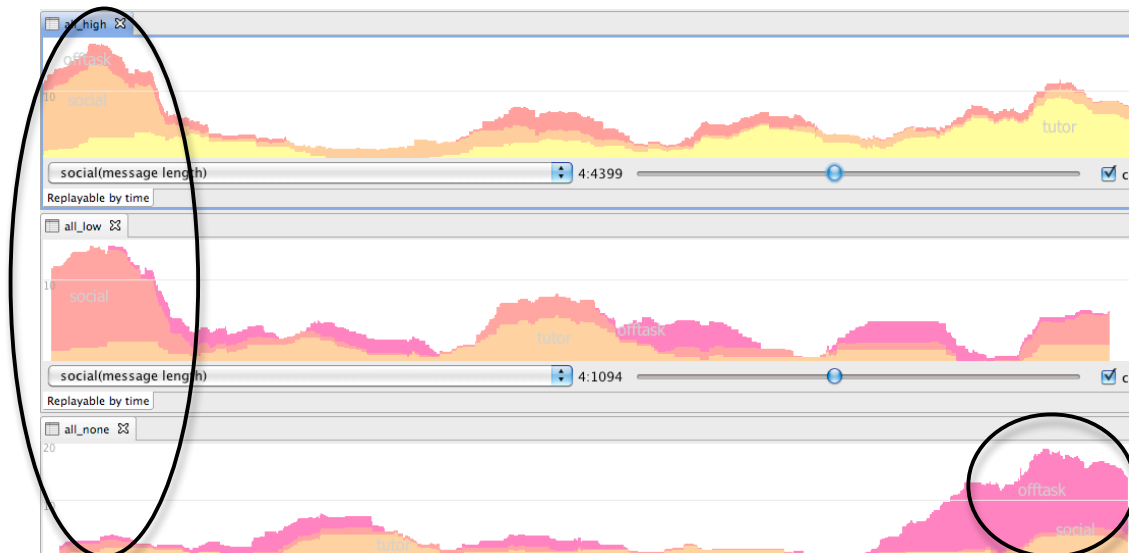
**Figure 1.** Sliding window visualization for windows of 1mn40s across all groups, showing distribution of features over time (horizontal axis). Units are in words per minute per group. Tutor (pastel) vs student (pink) is shown at top. Student turns are decomposed (bottom) into their component KCD and non-KCD sections.

## Partitioning groups

Armed with the knowledge of what a normal group (or subset of groups) would look like, we partitioned the groups according to the experimental conditions, contrasting the graphs for each. In Figure 2, for example, we see that the Power condition produced less talk in the (circled) discussion immediately after the Fuel KCD than the Neutral or Green conditions. Examining the transcripts in more detail, this appears to be because the introduction of the discussion related to Fuel, particularly when combined with the environmental bias (which discourages the use of fuels which produce a lot of waste heat – and increase power) generated a lot of rediscussing of values that might not otherwise have been reconsidered.



**Figure 2.** Sliding window visualizations for windows of 1mn40s partitioned by Goal condition (Top: green; Middle: power; Bottom: neutral). The curves are strikingly similar across conditions, apart from the discussion (green, circled) immediately after the fuel KCD (orange), which is noticeably lower in the power condition.



**Figure 3.** Sliding window visualization (window of 4mn43) for offtask, social and tutor abuse, portioned by Social condition (Top: high; Middle: low; Bottom: none). We see high social in the beginning for the social conditions, and high offtask at the end for the none condition.

We were particularly able to refine our understanding of the interaction between social conditions and offtask, social talk and tutor abuse. First, as is apparent in Figure 3, the majority of social talk happened during the beginning of the session in the High and Low social conditions and was almost completely absent in the None social condition. Upon examination of the transcripts the reason was immediately apparent: in the social

conditions, the tutor suggested that participants introduce themselves – which they then did, almost consistently in all groups. Tutor abuse can be seen to be steadily rising in the high condition, while being much more varied in the others. This shows the growing frustration that the students in the high social condition experienced with the tutor, which was performing too many social turns to remain a credible participant in the discussion. Last, the offtask towards the end is particularly marked in the end of the None condition, possibly as a reaction to the tutor leaving, presumably leading the students to believe that the task was over.

### **Identifying new variables for statistical analysis**

The differences between the Green and Power conditions, along with other, less noticeable ones when we partitioned both by learning outcome (post-test, adjusted to account for pretest values) and by condition lead us to believe that separating participation into individual discussion portions (the various KCD and non-KCD phases) would produce productive variables for examining correlation with learning. It turns out that in the Green condition, the discussion about Tmax was significantly greater (in total number of words) for high learners than for low learners,  $F(2,43) = 3.98$ ,  $p < .05$ , effect size 1.47 standard deviations, whereas there was no difference within the Power and Neutral conditions. In the Power condition, however, it was the discussion about Pmin which was greater for the high learners than for the low learners,  $F(2,43) = 3.65$ ,  $p < .05$ , effect size .43 standard deviations. Interestingly, in both cases, these particular subjects were the more cognitively demanding with their respective biases (Tmax should go as high as possible to get greater power, but needs compromise to get environmental friendliness; Pmin affects cycle efficiency and so is the only one of the four parameters for which there isn't a tradeoff between power and environmental friendliness).

Examining the direct relationship between the various social conditions and the learning outcome did not reveal any immediate answers. When combining this with information about offtask, however, a clearer picture can be drawn. We examined the amount of offtask during each portion of the discussion and its correlation with learning outcomes. Low learners had significantly more offtask during the KCD portion of the discussion than high learners,  $F(1,36) = 10.3$ ,  $p < .005$ , effect size 1.62 standard deviations. On the other hand, with respect to the Power and Green conditions there was no difference between higher and lower performing students regarding how much tutor abuse they uttered, but surprisingly, within the Neutral condition, higher performing students did more tutor abuse than lower performing students,  $F(2,43) = 3.87$ ,  $p < .05$ , effect size 1 standard deviation. This dual effect between social condition and on/off task behavior and between on/off task behavior and learning outcome illustrates how intricate such interactions between variables can be, particularly when time is factored back into the analysis.

To summarize, our new analysis shows that there is a distinct fluctuating pattern that can be expected of all the conversations scripted by this conversational agent. It shows that the effect of tutor social support on student social interaction is mostly a product of the introductory phase of the interaction (although it does extend to some extent over the whole interaction). It suggests an interaction between social condition, amount of offtask, and learning outcome, indicating that certain points of the interaction are more important (during the KCDs and during the final discussion) than others and that appropriate agent scaffolding might attempt to divert offtask behavior from happening at these points, or that offtask at this behavior might be diagnostic of groups with low learning performance. Finally, we identified a previously unsuspected interaction between participation during specific KCDs, respectively in the Green and Power conditions, and learning outcomes. Again this relationship can currently only be suggested as a predictor of learning, and not a cause of learning. Further studies might attempt to scaffold the interaction to favor patterns which have, in the past, been indicative of high learning and attempt to discern a causal relationship.

### **Discussion**

In this example, we show how interactive sliding windows can be integrated in different ways into both descriptive and quantitative methodologies, particularly when focusing on a search for irregularities. This example has certain properties which render it particularly suited to the methodology we are developing: it is strongly scripted over time, making it easy and feasible to cumulate groups together and to partition data into distinct sets of groups. Furthermore, there is a strong homogeneity between groups (each has two students, with similar goals) and an experimental design which provides a number of randomly assigned existing conditions. The data in this example was only coded in a limited way to show offtask and ontask and all visualizations were calculated using turn word count as a proxy for participation. More complex coding might have revealed more pertinent results, but it would then have been harder to attribute new findings to the introduction of sliding window visualizations as opposed to the new codings.

Interactive sliding windows can intervene at different times within existing methodologies. For an initial data exploration phase, they provide an interesting lens for understanding how a process develops over time, drawing a continuous link between session-wide summary statistics and properties of individual turns, and allowing the analyst to delve down into the raw data at any time. They also provide a way to examine what regularities and distinctions that exist in different conditions or sub-populations. This helped us show both than

some of the results of the initial analysis were less interesting than they appeared to be (the effect of social condition on amount of social participation was most marked in the introductions) and that others were more interesting (the interaction between both goal and social conditions and offtask and tutor abusing participation is particularly complex). On their own however, these visualizations are not suitable as a quantitative tool. As in our example, they are useful for identifying new variables that might potentially be used to show significant quantitative results. We did not report any abortive analytic paths, but the visualizations frequently helped us to cheaply explore different avenues in which quantification might have yielded results, but for which it was immediately apparent that there would be nothing significant to report. For example, we also used sub-population partitioning on post-hoc values such as the learning outcome. This showed an apparent linear relation between participation in the final discussion and learning, with its value being smallest in the low learners, largest in the high learners and in between in medium learners. This relationship turned out to be artifacts of the data partitioning, rather than a significant result. Partitioning by learning outcome *and* condition, however produced the significant result reported above. Finally, the dynamic nature of these visualizations, helps overcome a major limitation of time-scale changes – that there is no reason that an arbitrarily sized window (either in time or number of turns) should have a constant relation with complex group processes.

The visualizations proved particularly challenging to read, frequently suggesting patterns which turned out (on closer examination of the data) to be artifacts caused by certain groups or by looking at a specific window size. One of the issues is that our indicators only report averages and not distributions informing us of variance, etc. This effect can be slightly mitigated by synchronizing across multiple visualizations, so that their timescale is constant, the window size currently under examination is identical and the units are coherent across window sizes. The units must also be coherent across subpopulations of different sizes. In order to meaningfully compare visualizations, we made each indicator be computed per group and per minute. Not only did synchronization from Tatiana help us transition between visualizations and raw data, it also helped understand the non-intuitive relationship between an indicator computed over a given window and the events making up that window. It is useful to consider each event has having an area of influence comprising multiple windows, and to consider each window as being the first, chronologically, to include certain events and to exclude others.

## Conclusion

In this paper, we presented interactive sliding window visualizations and showed the role they can play in integrating with existing methodologies to explore data, produce more refined results and new insights into the irregular temporal processes which occur during collaborative learning. These visualizations address some of the granularity issues found in previous work attempting to quantify processes over time. They also draw upon work in visualization and exploratory sequential data analysis providing a method which fits nicely in the gap between descriptive and quantitative methodologies, allowing transitions between event properties and global indicators, and providing the means both to dive into the raw data and suggest new variables for quantification, fundamentally looking for regularity across populations and irregularities between populations and over time. We showed how these visualizations provided new insights on a previously analyzed corpus, showing that certain results were more interesting than they seemed, some less, and providing additional entirely new insights. We then discussed the limitations of these visualizations, highlighting that they are not intended to stand on their own, but to be used as a means of identifying potential irregularities which can then be explained qualitatively or confirmed quantitatively as related to learning and developmental processes. These visualizations also prove particularly appropriate when the data is scripted, providing a link between the intended script and the actual path through the script. Future work should investigate how these visualizations can be used for larger timeframes and in natural, rather than quasi-experimental, and non-educational settings.

We chose to focus on very superficial features of the conversations without using complex coding. This was in part in order to show the benefits over an existing analysis using the same features, but also because we see this as fitting into two wider strains of research. The first is the problem of data analysis where researchers attempt to find anchor points which allow them to explore a corpus which has not yet been analyzed or coded – sliding window visualizations assist in the identification of salient features which may be worthy of further interest. The second is the problem of providing dynamic support for collaborative learning. In our research into conversational agents, we have been hampered by the fact that agents do not have a deep understanding of how conversation unfolds over time. In a first step, we would like to identify temporal features which allow a conversational agent to gain insight into how well or poorly a group is performing. Once we have identified such features (and that they correlate with positive outcomes) we will be able to address the more complex issue of whether the agent is able to influence the interaction in a positive way, or in other words to examine whether we can also posit a causal effect between some kinds of features the agent could influence and a positive outcome.

## Acknowledgments

This work was funded by NSF SBE-0836012.

## References

- Ai, H., Kumar, K., Nguyen, D., Nagasunder, A., and Rosé, C. 2010. Exploring the Effectiveness of Social Capabilities and Goal Alignment in Computer Supported Collaborative Learning. *Proceedings of 10th International Conferences on Intelligent Tutoring Systems (ITS)*.
- Chiu, M. M., & Khoo, L. (2005). A new method for analyzing sequential processes: Dynamic multi-level analysis. *Small Group Research*, 36, 600-631.
- De Wever, B., Schellens, T., Valcke, M., & Van Keer, H. (2006). Content analysis schemes to analyse transcripts of online asynchronous discussion groups: A review. *Computers & Education*, 46(1) 6-28.
- Dyke, G., Adamson, D., Howley, I., Rosé, C. (2012) Towards Academically Productive Talk Supported by Conversational Agents. *ITS 2012*.
- Dyke, G., Lund, K., & Girardot, J.-J. (2010). Tatiana : un environnement d'aide à l'analyse de traces d'interactions humaines. *Technique et Science Informatiques*, 29(10), pp.1179-1205.
- Dyke, G., Lund, K., & Girardot, J.-J. (2009). Tatiana: an environment to support the CSCL analysis process. *CSCL 2009*, Rhodes, Greece, 58-67.
- Hmelo-Silver, C. E., Chernobilsky, E., & Nagarajan, A. (2009). Two sides of the coin: Multiple perspectives on collaborative knowledge construction in online problem-based learning. In K. Kumpulainen, C. E., Hmelo-Silver, & M. Cesar (Eds.). *Investigating Classroom interaction: Methodologies in action*. Sense Publishers.
- Jadallah, M., Anderson, R.C., Nguyen-Jahiel, K., Miller, M., Kim, I.-H., Kuo, L.-J., Wu, X., & Dong, T. (2011). Influence of a teacher's scaffolding moves during child-led small-group discussions. *American Educational Research Journal*, 48(1), 194-230.
- Jain, M., McDonough, J., Gweon, G., Raj, B., & Rose, C (2012). An Unsupervised Dynamic Bayesian Network Approach to Measuring Speech Style Accommodation. *EACL 2012*.
- Jeong, A. (2008). Sequentially analyzing and mapping the interactional processes of knowledge construction in online learning. *AERA 2008*, New York.
- Jermann, P., Mullins, D., Nuessli, M.-A., Dillenbourg, P. (2001). Collaborative Gaze Footprints: Correlates of Interaction Quality. In *Proceedings of CSCL 2011*. Hong Kong.
- Kapur, M. (2011) Temporality matters: Advancing a method for analyzing problem-solving processes in a computer-supported collaborative environment. *ijcscl* 6 (1), pp. 39-56.
- Kumar, R., Ai, H., Beuth, J., Rosé, C. P. (2010) Socially-capable Conversational Tutors can be Effective in Collaborative-Learning situations. *Intelligent Tutoring Systems (ITS) 2010*, Pittsburgh, PA
- Maldonado, R. M., Yacef, K., Kay, J., Kharrufa, A., & Al-Qaraghuli, A. (2011). Analysing frequent sequential patterns of collaborative learning activity around an interactive tabletop. *Proceedings of EDM 2011*.
- Mayfield, E., Rudnicky, A., and Rose, C. (2012). Computational Representation of Discourse Practices Across Populations in Task-based Dialogue. *ICIC 2012*.
- McGrath, J. E., & Tschan, F. (2004). *Temporal matters in social psychology: Examining the role of time in the lives of groups and individuals*. Washington: American Psychological Association.
- Muukkonen, H., Hakkarainen, K., Konsonen, K., Jalonen, S., Heikkil, A., Lonka, K., et al. (2007). Process-and context-sensitive research on academic knowledge practices: Developing CASS-tools and methods. In *Proceedings of CSCL 2007* (pp. 541–543). New Brunswick: ISLS.
- Reimann, P. (2009) Time is precious: Variable- and event-centred approaches to process analysis in CSCL research. *ijcscl* 4 (3), pp. 239-257.
- Reimann, P., Frerejean, J., & Thompson, K. (2009). Using process mining to identify decision making processes in virtual teams. Paper presented at *CSCL2009*. Rhodes/Greece.
- Sanderson, P. M., & Fisher, C. (1994). Exploratory sequential data analysis: Foundations. *Human-Computer Interaction*, 9(3/4), 251–317.
- Suthers, D. D. (2006) Technology affordances for intersubjective meaning making: A research agenda for CSCL. *ijcscl* 1 (3).
- Suthers, D. D., Dwyer, N., Medina, R., & Vatrappu, R. (2010). A framework for conceptualizing, representing, and analyzing distributed interaction. *ijcscl* 5(1), 5-42.
- Suthers, D. D., Teplovs, C. De Laat, M., Oshima, J., Zeini S. (2011). Connecting Levels of Learning in Networked Communities. *Workshop at CSCL 2011*.
- Thompson, K., Kennedy-Clark, S., Markauskaite, L., & Southavilay, V. (2011). Capturing and analysing the processes and patterns of learning in collaborative learning environments. *CSCL 2011*. Hong-Kong.
- Wang, H., & Rosé, C. P. (2007). A Process Analysis of Idea Generation and Failure. *Proceedings of the Annual Meeting of the Cognitive Science Society*.
- Wise, A. F., & Chiu, M. M. (2011). Analyzing temporal patterns of knowledge construction in a role-based online discussion. *ijcscl* 6 (3), pp. 445-470.