Sensitivities to Early Exchange in Synchronous Computer-supported Collaborative Learning (CSCL) Groups

Manu Kapur
National Institute of Education, Singapore, manu.kapur@nie.edu.sg

John Voiklis[1], **Charles K. Kinzer**[2]
Teachers College, Columbia University, New York, USA
[1]jv37@tc.edu, [2]kinzer@tc.edu

Abstract: This study reports the impact of high sensitivity to early exchange in 11th-grade, CSCL triads solving well- and ill-structured problems in Newtonian Kinematics. Analysis of the evolution of participation inequity (PI) in group discussions suggested that participation levels tended to get locked-in relatively early on in the discussion. Similarly, high (low) quality member contributions made earlier in a discussion did more good (harm) than those made later on. Both PI and differential impact of member contributions suggest a high sensitivity to early exchange; both significantly predicting the eventual group performance, as measured by solution quality. Consequently, eventual group performance could be predicted based on what happened in the first 30-40% of a discussion. In addition to theoretical and methodological implications, implications for scaffolding CSCL groups are drawn.

Introduction

The role of collaborative interaction and participation is central to the socio-constructivist perspective on learning; a perspective that undergirds much of computer-supported collaborative learning (CSCL) research (Scardamalia & Bereiter, 2003; Stahl, 2005). Naturally, a major concern in CSCL research is why some groups are more successful than others. Historically, researchers have sought to address this concern by focusing on the effects of pre-existing group characteristics and member traits (e.g., group size, group ability, prior knowledge, heterogeneity, status, personality composition, learning styles, etc.) on group performance (e.g., Webster & Driskell, 1983; McAuliffe, 1991; Webb & Kenderski, 1984; Cohen, 1982; Webb, 1984; Sharan & Shachar, 1988; Lam, 1997).

Lately, though, there has been a push towards unpacking group processes, in particular the complexities of interactional dynamics and how it influences group performance; the nature of member interactions and participatory patterns forming key objects of inquiry. In fact, a realization of the inherent complexity of interactional dynamics is giving way to a more *temporal* and *emergent* view of how groups function and perform (Stahl, 2005; Arrow, McGrath, & Berdahl, 2000). This presents a unique challenge to traditional analytical measures and methods for analyzing group processes (Barab, Hay, Yamagata-Lynch, 2001); existing methods continue to take cumulative accounts of member interactions (e.g., categorization of interactional content, rating of discussion quality, member perceptions, and so on) and relate them to group performance. While these accounts are useful, they fail to fully utilize the temporal, evolutionary information embedded in the data (Kapur, Voiklis, Kinzer, & Black, 2006). Therein lies the need for the research reported in this paper.

We start with a brief review of how interaction and participation have been studied in CSCL research. A substantial amount of literature attempts to understand group processes using qualitative analytical methods, which provide insightful and meaningful micro-genetic accounts. For the present purposes, however, our analysis is limited to quantitative approaches, typically involving quantitative content analysis (QCA) (Chi, 1997) of interactional data; the use of QCA or what is commonly also referred to as "coding and counting" being pervasive in examining the nature of interaction and participation in CSCL research (Rourke & Anderson, 2004).

Nature of Interactions and Participation in CSCL Groups

Nature of Interactions

Because learners interact with and influence each other in the process of problem solving, these interactions form the most important units of analyses for research; problem-solving interactions have been used extensively in investigating productivity conditions of small, collaborative groups (e.g., Scardamalia, 1989, 1992; Scardamalia &

Bereiter, 1993; Bereiter, Scardamalia, Cassells, & Hewitt, 1997; Jonassen & Kwon, 2001; Barron, 2003; Cohen et al., 2002; Schellens, Van Keer, & Valcke, 2004; Zumbach, Schonemann, & Reimann, 2005; Spada, Meier, Rummel, & Hauser, 2005; Lee, Chan, & van Aalst, 2006; Kapur, 2006).

A common analytical thread runs through these investigations: they typically employ one or more coding/rating schemes, which when applied to the interaction data result in a cumulative frequency or relative frequency distribution of interactions across the categories of the coding/rating scheme (e.g., depth of explanations, functional content of interactions, misconceptions, quality, correctness, etc.). These distributions essentially tally the amount, proportion, and type of interactions vis-à-vis the interactional coding/rating scheme (Suthers, 2006). Significant links are then sought between quantitatively-coded interactional data and outcomes, such as quality of group performance and group-to-individual transfer (see Rourke and Anderson (2004) for a discussion on the validity of QCA).

Notwithstanding the empirically-supported significant links between the nature of interactional activity and group performance, interpreting findings from interactional coding/rating schemes is limited by the very nature of the information tapped by these measures. For example, what does it mean if a group discussion has a high proportion of, say, problem definition type of interactional activity? Clearly, answering this question is problematic when one considers two contrasting possibilities, both corresponding to the same proportion of problem definition type of interactional activity in a group. It could be that the group's attempt at defining the problem was spread throughout the discussion, or perhaps all such activity was clustered together in a coherent phase during the discussion. Furthermore, problem definition contributions that are temporally far apart in a discussion carry the same weight in the cumulative count or proportion; one that comes later in a discussion is given the same weight as one that comes earlier. Such an analysis, while informative, does not take the temporality of interactions into account, i.e., the time order of interactions in the problem-solving process. In light of the complexities of interactional dynamics in CSCL, it is surprising how frequently this assumption of *temporal homogeneity* is made without justification or validation. This study was designed to examine this assumption.

Participation

With regard to participatory patterns of group members, previous research has attempted to link individual participation with group performance as well as subsequent group-to-individual transfer. The study of participation in collaborative settings has primarily been studied at the individual level. Typically, research has focused on questions like - how an individual's participation rate in a group affects his/her achievement gains, or, how being part of a group with intensive interactional activity affects an individual's achievement gains, and other variants of the same (Cohen, 1994; Schellens, Van Keer, Valcke, & De Wever, 2005). Hence, they may suffer from ecological and atomistic fallacies in moving back and forth between the interpretation of findings at the individual and group levels.

An obvious work-around is to consider participation inequity as a group-level construct, as operationalized in the study described herein. After all, high participation inequity implies that group performance is primarily influenced by one or two dominant members. This leaves little opportunity for multiple perspectives, strategies, and solutions to be shared and discussed. Yet, the effect of participation inequity (as a group-level construct) on group performance remains relatively unexplored (e.g., Kapur & Kinzer, 2005).

In addition to focusing primarily on the effects of an individual's participation, previous research on participation in CSCL also makes a temporal homogeneity assumption: participation rates are cumulatively summed over the entire discussion (e.g., Schellens et al., 2005; Kapur & Kinzer, in press). By cumulative summing, however, one does not know whether member participation was greater earlier in the discussion or later; whether participation inequities evolve slowly or quickly. In other words, two things are ignored or assumed to be negligible in non-weighted summing of participation instances: temporal variation in a given member's participation over the course of a discussion, and consequently, participation inequity at the group level. Once again, there is an assumption of *temporal homogeneity*, this time for group participation inequity. However, this assumption is made without justification or validation.

Purpose

The purpose of this study was to test the temporal homogeneity assumption within the short-term, problem-solving efforts of synchronous, CSCL groups. We started by examining the temporal homogeneity of participatory patterns, specifically the evolution of participation inequity in CSCL groups. Our analysis revealed that the assumption of temporal homogeneity did not hold. Group performance was highly sensitive to participation inequities in the early exchange between group members; inequities during this sensitive period of early exchange seemed to get “locked in” for the rest of the discussion. In turn, this led us to examine the assumption of temporal homogeneity in terms of the impact that the quality of member interactions had on a group’s discussion.

Method

Research Context and Data Collection

Participants included sixty 11th-grade students (46 male, 14 female; 16-17 years old) from the science stream of a co-educational, English-medium high school in Ghaziabad, India. They were randomized into 20 triads and instructed to solve either a well-structured (WS) or an ill-structured (IS) authentic car accident scenario that required the application of concepts in Newtonian kinematics. The study was carried out as part of their regular classroom activity, where group members communicated with one another only through synchronous, text-only chat. The 20 automatically-archived transcripts, one for each group, contained the group discussions as well as their solutions, and formed the data used in our analyses.

Procedure

A well- and an ill-structured problem scenario were developed consistent with Jonassen’s design theory typology for problems (2000). Both problem scenarios dealt with a car accident scenario that required participants to apply principles of Newtonian Kinematics and Laws of Friction to solve them (see Appendix). Content validation of the two problem scenarios was achieved with the help of two physics teachers from the school with experience in teaching those subjects at the senior secondary levels. The teachers also assessed the time students needed to solve the problems. Feedback from the teachers resulted in minor modifications to the problem scenarios, which were them deemed to be consistent with the school's curriculum.

Problem classification validation was then undertaken by having the top three tenth-grade science students and the two teachers classify the two problems as being either well- or ill-structured. All students’ and teachers’ classifications were unanimously consistent with those of the researchers. The same three students were also asked to solve the problems to confirm that two hours would be sufficient time for the task. All of them completed the problems and submitted their work in about one hour. However, for group work, we decided to give each group two hours to allow sufficient time for group interaction and discussion; naturally, we didn’t want a lack of time to be a confounding factor. Ultimately, all groups completed the problem in the allotted time.

The study was carried out in the school’s computer laboratory where participants normally engage in a substantial amount of curricular problem solving activities. The online synchronous collaborative environment was a java-based, text-only chat application running on the Internet. Despite these participants being technologically savvy in using online chat, they were familiarized in the use of the synchronous text-only chat application prior to the study. Group members could only interact within their group. Each group’s discussion and solution were automatically archived as a text file to be used for analysis. A seating arrangement ensured that participants of a given group were not proximally located so that the only means of communication between group members was synchronous, text-only chat. To mitigate status effects, we ensured that participants were not cognizant of their group members’ identities; the chat environment was configured so that each participant was identifiable only by an alpha-numeric code. Cross-checking the transcripts of their interactions revealed that participants followed the instruction not to use their names and instead used the codes when referring to each other. No help regarding the problem-solving task was given to any participant or group during the study. Furthermore, no external member roles or division of labor were suggested to any of the groups. The procedures described above were identical for both WS and IS groups. The time stamp in the chat environment indicated that all groups made full use of the allotted time of two hours and solved their respective problems.

Results

Evolution of Participation Inequity (PI)

In this study, PI was conceived as a group-level construct and operationalized as the standard deviation (SD) of the three member participation proportions (number of utterances by a member as a proportion of total utterances) within each group. A low SD implies closely clustered participation ratios within a group, i.e., a participation pattern that is more or less uniform and equitable. For example, the *SD* of the participation proportions

.4, .3, and .3 equals .06. Thus, a low *SD* implies closely clustered participation proportions within a group—an *equitable* participation pattern. On the other hand, a high SD implies a discussion that is dominated by one or two members within the group, i.e., an inequity in the participation of members in the group. For example, the *SD* of participation proportions .8, .15, and .05 equals .41. Thus, a high *SD* implied a discussion dominated by one or two members within the group—an *inequitable* participation pattern. Next, PIs after each utterance in a discussion were calculated, giving the level of PI in the discussion up to any given utterance. Plotting these values over time (defined notionally with utterances as *ticks* on the evolutionary clock) reflected the temporal development of PI for the 20 groups. Figure 1 shows the typical trajectories for WS and IS groups.

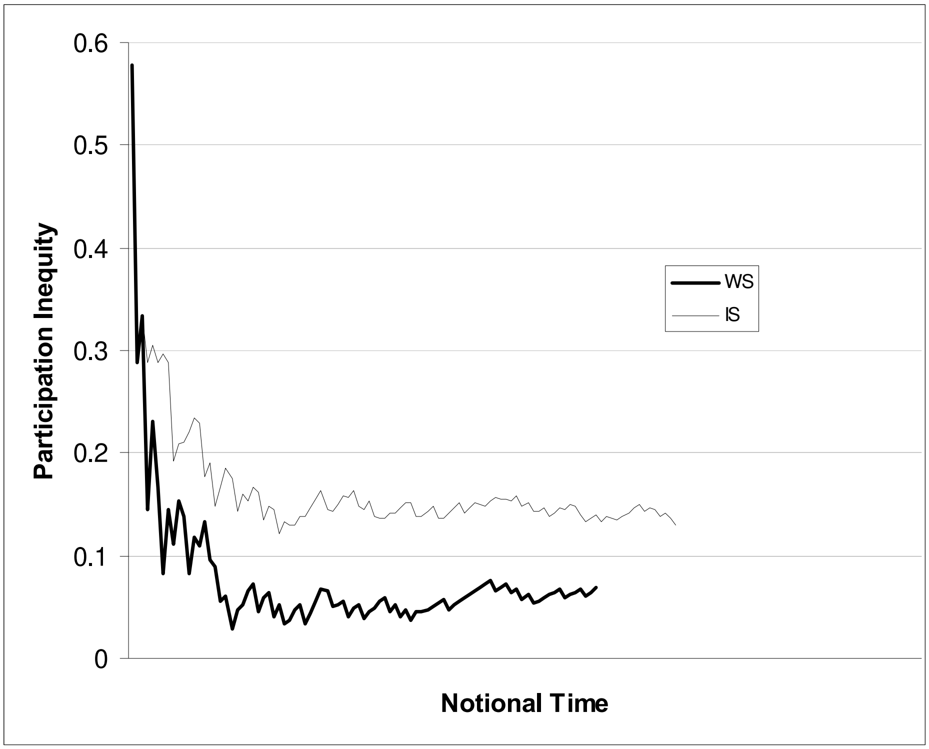

Figure 1. Evolution of participation inequities across problem types

What was surprising was how sensitive the evolution of these trajectories was to the early exchange between group members in both WS and IS groups. This can be seen in the sharp fluctuations in the trajectories in Figure 1 before they quickly settled into an inequity plateau. The main difference between WS and IS groups seemed to be that the former typically settled into a lower plateau (i.e., lower PI; higher equity) whereas the latter into a higher plateau (i.e., higher PI; lower equity). This difference between WS and IS groups is interesting in and of itself and we unpack it in greater detail elsewhere (Kapur & Kinzer, in press). For the present purposes, however, we focus on the patterns across the groups, treating problem type as a control variable. This would allow us to focus on what is significant across the WS and IS groups, i.e., PI evolution was not a gradual process but one that was highly sensitive to early exchange. Critically, inequities during this sensitive period of early exchange seemed to get "locked in" for the rest of the discussion, i.e., once settled into an inequity plateau early in the discussion groups were, on average, unlikely to escape it.

Given the above finding, the logical question becomes: how does the early lock-in of PI influence group performance? To answer this question, group performance was operationalized as the quality of group solution, independently rated by two doctoral students on a 9-point rating scale (Table 1) with an inter-rater reliability (*Krippendorff's alpha*) of .95. An analysis of variance (ANOVA) (Table 2) showed that, controlling for problem type, PI was a significant predictor of eventual group performance (F = 8.484, p = .010); High PI resulting in low group performance, on average. Levene's test for equality of error variance was not significant (F = .782, p = .388).

Table 1. *Rubric for coding quality of group solution*

Quality	Description
0	Solution is weakly supported, if at all
1	Solution supported in a limited way relying either on a purely quantitative or a qualitative argument with little, if any, discussion and justification of the assumptions made
2	Solution is only partially supported by a mix of both qualitative and quantitative arguments; assumptions made are not mentioned, adequately discussed, or justified to support the decision
3	Solution synthesizes both qualitative and quantitative arguments; assumptions made are not adequately discussed and justified to support the decision
4	Solution synthesizes both qualitative and quantitative arguments; assumptions made are adequately discussed and justified to support the decision

Note: Mid-point scores of .5, 1.5, 2.5, and 3.5 were assigned when the quality of solution was assessed to be between the major units 0, 1, 2, 3, and 4, making the scale essentially a 9-point scale.

Table 2. *Model summary for predicting group performance from PI*

Source	***SS***	***df***	***MS***	***F***	***p***	**Partial Eta Squared**	**Power**[a]
Intercept	25.690	1	25.690	31.743	.000	.651	1.000
Participation Inequity	6.867	1	6.867	8.484	.010	.333	.784
Problem Type	.434	1	.434	.537	.474	.031	.106
Error	13.758	17	.809				
Total	81.250	20					

[a] Computed using alpha = .05
[b] R Squared = .516 (Adjusted R Squared = .459)

These findings support the view that early lock-in of PI is significant, for it suggests that the impact of early exchange on group performance is much greater than what comes later. The seeds of eventual group performance seem to be sown fairly early in a group's discussion. Because a lock-in of participation levels also implies a lock-in to the dominant members' proposals or contributions, it meant that one could no longer assume member contributions made a temporally homogenous impact on group performance. It is this analysis that we turn our attention to next.

Evolution of Differential Impact of Member Contributions

Quantitative content analysis (QCA; Chi, 1997) was used to segment utterances into one or more interaction units. The interaction unit of analysis was semantically defined as the impact(s) that an utterance had on the group discussion (Bransford & Nitsch, 1978). Adopting Kapur et al.'s methodology (2006), an impact value of 1, -1, or 0 was assigned to each interaction unit depending upon whether it moved the group discussion toward (impact = 1) or away (impact = -1) from the goal of the activity—a solution state of the given problem, or maintained the status quo (impact = 0). Therefore, each discussion was reduced to a temporal string of 1s, -1s, and 0s, i.e., a non-random walk (Ross, 1996). Two trained doctoral students independently coded the interactions with an inter-rater reliability (*Krippendorff's alpha)* of .85. This impact coding is illustrated through a brief micro-analytical analysis of the following excerpt containing an exchange between group members S1 and S2.

S1 > *are we going to apply frictional retardation for the reaction time also?*	-1
S2 > *no, because reaction time is the time between watching the boy and applying the brakes so in this time car must be accelerating*	1, 1
S1 > *but I think we must not forget that the car is moving on the same road on which the incident occurs and the road is providing the retardation*	-1, -1
S2 > *but maximum speed is at the instant when he applied the brake*	1
S1 > *but earlier you said that the car will accelerate after perceiving the boy*	-1
S2 > *I said so because his foot must be on accelerator during reaction time*	1
S1 > *Now I understand… please proceed to write your answer*	1, 1

Recall that the problem involved a car-accident scenario. In this excerpt, S1 and S2 are trying to decide whether or not reaction time of the driver of the car that was involved in the accident should factor into their calculations. The excerpt starts with S1 asking a question about applying frictional retardation during reaction time of the driver. Being a misconception, it was rated as having a negative impact (-1) on the group's progress towards solving the problem—a collective divergence from the goal of solving the problem. S2 evaluates S1's question and says '*no*,' attempting to correct the misconception. Hence, its positive (+1) impact rating – a local divergence between S1 and S2 but a collective convergence towards the goal. In the same utterance, S2 elaborates why frictional retardation should not to be applied, further positively impacting the group's progress – continued local divergence but increasing collective convergence towards the goal. The argument continues with S1 persisting with the misconception (assigned negative impacts) until S2 is able to convince S1 otherwise (assigned positive impacts), thereby converging on a correct understanding of this aspect (dealing with friction during reaction time) of the problem given to them – both local as well as collective convergence toward the goal is achieved. Note that had S2 wrongly evaluated and agreed to S1's misconception, S2's impact ratings would have been negative, which, without any further correction, would have led the group to diverge from a correct understanding of the aspect of the problem being considered.

As applied, however, the impact coding not only takes into account local convergence but also convergence of the group as a whole toward solving the problem; impact ratings are meaningful only in relation to preceding utterances (Bransford & Nitsch, 1978) and take into account the sequence and temporality of collaborative interactions (Kapur et al., 2006). Other examples of highly convergent discussion episodes would include agreement with and positive evaluation and development of correct understandings of the problem, solution proposals, and problem solving strategies.

More formally, let n_1, n_{-1}, and n_0 denote the number of interaction units assigned the impact values 1, -1, and 0 respectively up to a certain utterance in a discussion. Then, up to that utterance, the mean impact (henceforth referred to as convergence) in terms of moving the group towards or away from the goal is given by the mean distance of the Markov walk, $C = \frac{n_1 - n_{-1}}{n_1 + n_{-1}}$ ($-1 < C < 1$). Convergence values were calculated after each utterance, resulting in a notional time series representing the evolution of member contributions' impact on the group discussion. Plotting the convergence value on the vertical axis and time (defined notionally with utterances as *ticks* on the evolutionary clock) on the horizontal axis, one gets a representation (also called a fitness curve) of the problem-solving process as it evolves in time. Figures 2 and 3 present the four major types of fitness curves that emerged from the discussion of the 20 problem-solving groups in our study. These four fitness curves contrast the high- with the low-performing groups across WS and IS groups.

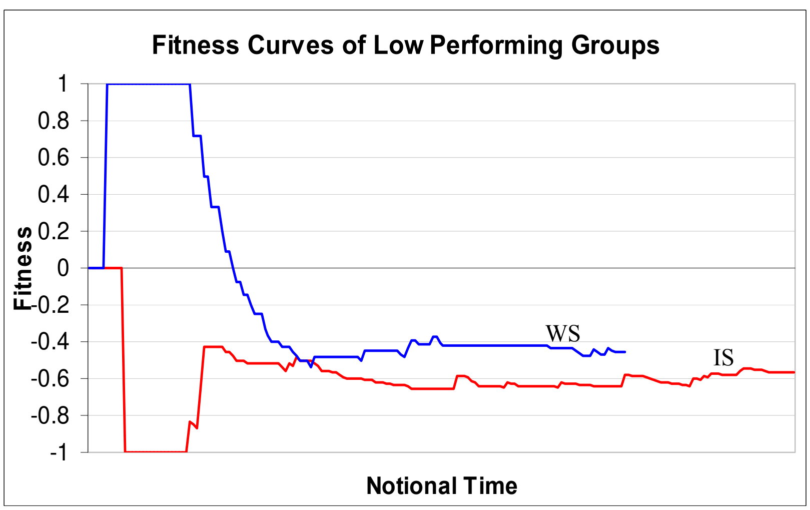

Figure 2. Fitness curves of low-performing groups across problem types

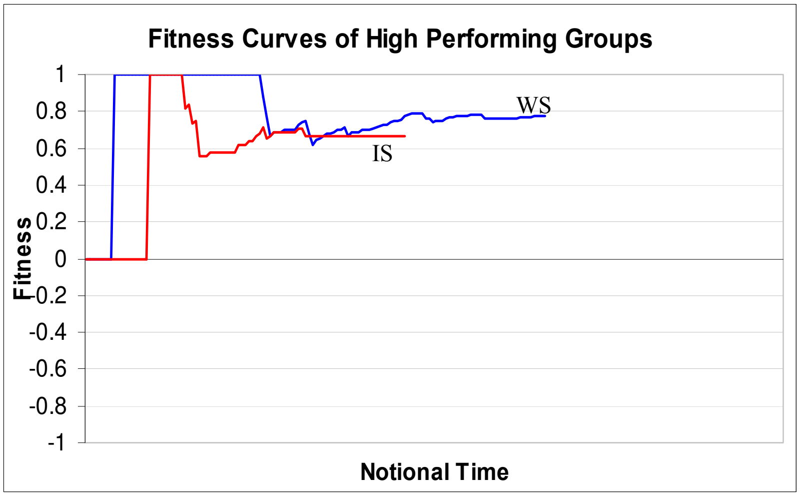

Figure 3. Fitness curves of high-performing groups across problem types

Interpreting Fitness Curves

It is easy to see that the convergence value always lies between -1 and 1. The closer the value is to 1, the higher the convergence, and the closer the group is to reaching a solution. The end-point of the fitness curve represents the final fitness level or convergence of the entire discussion. From this, the extent to which a group was successful in solving the problem can be deduced. Furthermore, one might imagine that an ideal fitness curve is one that has all the moves or steps in the positive direction, i.e., a horizontal straight line with fitness equaling 1. However, the data suggests that, in reality, some level of divergence of ideas may in fact be a good thing (Schultz-Hardt et al., 2002; Kapur, 2006), as can be seen in the fitness curves of both the high-performing groups.

The shape of the fitness curve, therefore, is also informative about the paths respective groups take toward problem solution. For example, in Figure 2, both the low-performing groups converged at approximately the same (negative) fitness levels, but their paths leading up to their final levels were quite different. The WS group showed a sharp fall after initially moving in the correct direction (indicated by high fitness initially). The IS group, on the other hand, tried to recover from an initial drop in fitness but was unsuccessful, ending up at approximately the same fitness level as the well-structured group. Further, comparing the high-performing groups (Figure 3) with the low-performing groups (Figure 2), one can see that the discussions of high-performing groups had fewer utterances, regardless of problem type. Finally, all fitness curves seemed to settle into a *fitness plateau* fairly quickly. Again, as with PI, once the fitness was "locked in", groups found it increasingly difficult to escape it.

What is most interesting is that this temporal, albeit descriptive, examination of fitness curves provides a view of paths to a solution that are lost in analysis systems that assume temporal homogeneity and consider only a given point in the solution process, thus assuming that similar behaviors or states at a given point are arrived at in similar ways. As different paths can lead to similar results, unidimensional analyses that cumulatively consider only single points in time (often only the solution state) are not consistent with what this study's data suggest about CSCL processes.

Most important is a mathematical property of convergence. Being a ratio, convergence is more sensitive to initial contributions, both positive and negative, than those made later in the process. This can be easily seen because with each positive (or negative) contribution, the ratio's numerator is increased (or decreased) by 1. However, the denominator in the ratio always increases, regardless of the contribution being positive or negative. Therefore, when a positive (negative) contribution comes earlier in the discussion, its impact on convergence is greater because a unit increment (decrement) in the numerator is accompanied by a denominator that is smaller earlier than later. Said another way, this conceptualization of convergence allows us to test the *differential temporal impact* hypothesis: *"good" contributions made earlier in a group discussion, on average, do more good than if they were made later. Similarly, "bad" ones, on average, do more harm if they come earlier than later in the discussion*. To test this hypothesis, the relationship between convergence and group performance was explored by running a temporal simulation on the data set.

Testing the Differential Temporal Impact Hypothesis

The purpose of the simulation was to determine if the level of convergence in group discussion provided an indication of the eventual group performance. Recall that group performance was operationalized as the quality of group solution. The discussions of all 20 groups were each segmented into 10 equal parts. For example, a discussion comprising 300 utterances was divided into 10 parts of 30 utterances each; a discussion comprising 150 utterances was divided into 10 parts of 15 utterances each, and so on. At each tenth, the convergence value up to that point was calculated. This resulted in 10 sets of 20 convergence values; the first set corresponding to convergence in the discussion after 10% of the discussion was over, the second after 20% of the discussion was over, and so on until the tenth set, which corresponded to the final convergence value of the discussion, i.e., after 100% of the discussion had occurred. A simulation was then carried out by regressing group performance on convergence values at each tenth of the discussion (hence, a temporal simulation), controlling for problem type (WS or IS) each time. The *p*-value corresponding to the statistical significance of the predictive power of convergence at each tenth of the discussion on eventual group performance was plotted on the vertical axis (see Figure 4). C1 through C10 denote the 10 equally-spaced instances in each discussion at which the convergence values were calculated. The simulation suggested that, on average, at some point after 30% but before 40% of the discussion is over (i.e., between C3 and C4 in Figure 4), the convergence value is able to predict eventual group performance at the .05 level of significance or better. This shows that convergence is a powerful measure that is able to model the impact that early contributions have on eventual group performance.

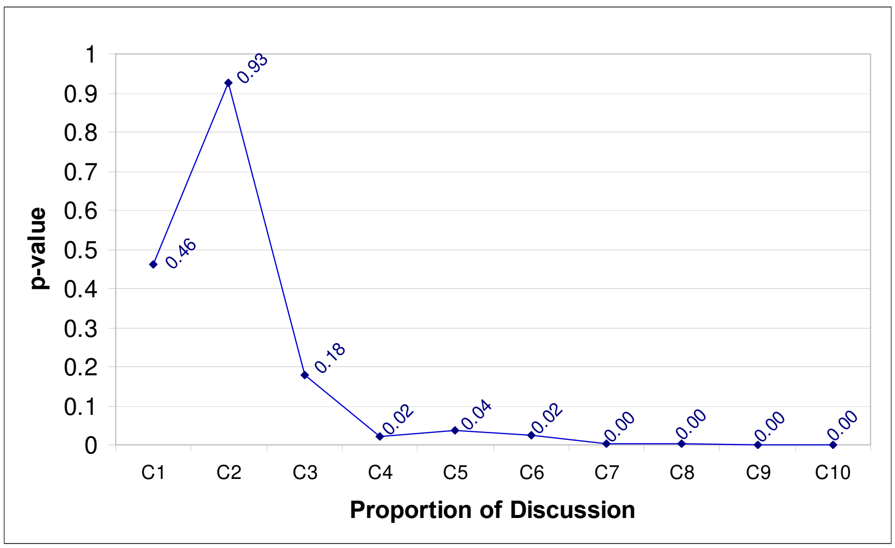

Figure 4. Simulation of the significance of convergence in predicting group performance

More importantly, this shows that the differential temporal impact hypothesis holds up to empirical scrutiny; eventual group performance is highly sensitive to early exchange, not just in terms of PI but also differential impact of member contributions. It is important to note that while both PI and convergence independently predicted group performance, they were also significantly negatively correlated (r = -.538, p = .014). Notwithstanding, when included in an ANOVA together (see Table 3), both remained significant predictors of group performance (F = 7.144, p = .017 & F = 16.789, p = .001 respectively), controlling for problem type.

Table 3. *Model summary for predicting group performance from PI and Convergence*

Source	*SS*	*df*	*MS*	*F*	*p*	Partial η^2	Power[a]
Intercept	.218	1	.218	.520	.481	.032	.104
Participation Inequity	2.998	1	2.998	7.144	.017	.309	.709
Convergence	7.045	1	7.045	16.789	.001	.512	.970
Problem Type	.156	1	.156	.371	.551	.023	.088
Error	6.714	16	.420				
Total	81.250	20					

[a] Computed using alpha = .05; [b] R Squared = .764 (Adjusted R Squared = .720)

Discussion

This study revealed novel insights into the problem-solving process of CSCL groups. Its most important implications stem from the finding that the sensitivity to early exchange tends to lock-in participation levels, which eventually determined how successful a group was in solving the given problem. This finding becomes even more significant because participation in high-performing groups is consistently a strong predictor of subsequent individual learning gains (e.g., see Barron, 2003; Cohen et al., 2002; Schellens et al., 2005). Furthermore, because a lock-in of participation levels also implies a lock-in to dominant members' proposals, those members' high (low) quality contributions have a greater positive (negative) impact on the eventual outcome when they come earlier than later in a discussion. Our findings indicate that eventual group performance could be predicted based on what happens in the first 30-40% of a discussion. This is not to say that contributions made later in a discussion are not important. Instead, once a discussion gets locked in, it gathers inertia, and it becomes increasingly difficult for individuals' subsequent proposals to make an impact proportional to their quality.

Implications for Scaffolding

If, as this study suggests, group performance is highly sensitive to the early exchange in the discussion, then this insight bears important implications for scaffolding synchronous, small-group, CSCL discussions to achieve optimal outcomes. For example, if one's interest is primarily in maximizing group performance, the insight suggests a need for scaffolding early in the discussion, since the impact of early interactional activity on eventual group performance seems to be greater. Scaffolding earlier parts of a group discussion may increase its likelihood of settling into lower inequity and higher fitness plateaus; better group performance is predicated by low and high inequity and fitness plateaus, respectively. This is also consistent with the notion of fading: having scaffolded the early exchange, the scaffolds can be faded. For example, instead of scaffolding the entire process of problem solving using process scaffolds, it may only be necessary to scaffold how a group analyzes and frames the problem, as these often occur early in problem-solving discussions. Such an approach stands somewhat in contrast with the common practice of blanket scaffolding of the CSCL processes (e.g., through the use of collaborative scripts). The above are testable hypotheses that emerge from this study and we invite the field to test and extend this line of inquiry.

Implications for Conceptualizing CSCL Groups

Interestingly, sensitivity to early exchange exhibited by CSCL groups in our study seems analogous to sensitivity to initial conditions exhibited by many complex adaptive systems (Bar-Yam, 2003; Arrow et al., 2000)); the idea being that small changes initially can lead to vastly different outcomes over time, which is what we found in our study. Furthermore, the locking-in mechanism is analogous to *attractors* in the phase space of complex systems (Bar-Yam, 2003). Phase space refers to the maximal set of states a complex system can possibly find itself in as it evolves. Evidently, a group discussion has an infinite phase space, yet the nature of early exchange can potentially determine whether it organizes into higher or lower inequity and fitness attractors. Thus, an important theoretical and methodological implication from this finding is that CSCL research needs to pay particular attention to the temporal aspects of interactional dynamics. As this study demonstrates, studying the evolution of interactional patterns can be insightful, presenting counterintuitive departures from assumptions of linearity in, and temporal homogeneity of, the problem solving process (Kapur et al., 2006).

Conclusion

This study was designed to examine the temporal homogeneity assumption that is often made in the study of CSCL groups. In particular, it demonstrated the inadequacy of assuming temporal homogeneity of participation patterns as well as the impact of member contributions. Our analysis revealed that group performance was highly sensitive to early interactional activity; both participation inequity and differential impact of member contributions significantly predicting the eventual group performance. All in all, this study offers preliminary yet compelling insights into the nature and dynamics of problem-solving CSCL groups. We fully acknowledge that our findings are technically bound by the context within which this study was conducted. Additionally, there are other factors such as prior knowledge, writing ability, gender, group composition, learning styles and dispositions, the nature of the task itself, affordances of the online chat environment, etc. that could just as well have influenced our findings. Still, in taking these first, essential steps toward understanding of how temporality affects CSCL group functions and performance, we call for further efforts within this line of inquiry.

References

Arrow, H., McGrath, J. E., & Berdahl, J. L. (2000). *Small groups as complex systems: Formation, coordination, development, and adaptation*. Thousand Oaks, CA: Sage Publications.

Barab, S. A., Hay, K. E., & Yamagata-Lynch, L. C. (2001). Constructing networks of action-relevant episodes: An in-situ research methodology. *The Journal of the Learning Sciences*, *10*(1&2), 63-112.

Barron, B. (2003). When smart groups fail. *The Journal of the Learning Sciences*, *12*(3), 307-359.

Bar-Yam, Y. (2003). *Dynamics of Complex Systems*. Reading, MA: Addison Wesley

Bereiter, C., Scardamalia, M., Cassells, C., & Hewitt J. (1997). Postmodernism, knowledge building, and elementary science. *The Elementary School Journal*, *97*(4), 329-340.

Bransford, J.D., & Nitsch, K. E. (1978). Coming to understand things we could not previously understand. In J. F. Kavanaugh and W. Strange (Eds.), *Speech and Language in the Laboratory, School, and Clinic*. Harvard, MA: MIT Press.

Chi, M. T. H. (1997). Quantifying qualitative analyses of verbal data: A practical guide. *The Journal of the Learning Sciences*, *6*(3), 271-315.

Cohen, E. G. (1982). Expectation states and interracial interaction in school settings. *Annual Review of Psychology*, *8*, 109-235.

Cohen, E. G. (1994). *Designing Groupwork: Strategies for Heterogeneous Classrooms*. (Eds.). New York: Teachers College Press.

Cohen, E. G., Lotan, R. A., Abram, P. L., Scarloss, B. A., & Schultz, S. E. (2002). Can groups learn? *Teachers College Record*, *104*(6), 1045-1068.

Jonassen, D. H. (2000). Towards a design theory of problem solving. *Educational Technology, Research and Development*, *48*(4), 63-85.

Jonassen, D. H. & Kwon, H. I. (2001). Communication patterns in computer-mediated vs. face-to-face group problem solving. *Educational Technology, Research and Development*, *49*(1), 35-52.

Kapur, M. (2006). Productive failure. In S. Barab, K. Hay, & D. Hickey (Eds.), *Proceedings of the International Conference on the Learning Sciences* (pp. 307-313). Mahwah, NJ: Erlbaum.

Kapur, M., & Kinzer, C. (in press). Examining the effect of problem type in a synchronous computer-supported collaborative learning (CSCL) environment. *Educational Technology, Research & Development*.

Kapur, M., & Kinzer, C. (2005, April). The effect of problem type on collaborative problem solving in a synchronous computer-mediated environment. Paper presented at the annual meeting of the American Educational Research Association, Montreal, Canada.

Kapur, M., Voiklis, J., Kinzer, C., & Black, J. (2006). Insights into the emergence of convergence in group discussions. In S. Barab, K. Hay, & D. Hickey (Eds.), *Proceedings of the International Conference on the Learning Sciences* (pp. 300-306). Mahwah, NJ: Erlbaum.

Lam, S. K. (1997). The effects of group decision support systems and task structures on group communication and decision quality. *Journal of Management Information Systems*, *13*(4), 193-215.

Lee, E. Y. C., Chan, C. K. K., & van Aalst, J. (2006). Student assessment of collaborative learning in a CSCL environment. *International Journal of Computer-Supported Collaborative Learning*, *1*(1), 57-87.

McAuliffe, T. (1991). Status rules of behavior in scenarios of cooperative learning groups. Paper presented at the Annual Meeting of the American Educational Research Association, Chicago.

Ross, S. M. (1996). *Stochastic processes*. New York: John Wiley & Sons, Inc.

Rourke, L., & Anderson, T. (2004). Validity in quantitative content analysis. *Educational Technology, Research and Development*, *52*(1), 5-18.

Scardamalia, M. (1989). Computer-supported intentional learning environments. *Journal of Educational Computing Research*, *5*(1), 51-68.

Scardamalia, M. (1992). Educational applications of a networked communal database. *Interactive Learning Environments*, *2*(1), 45-71.

Scardamalia, M., & Bereiter, C. (2003). Knowledge Building. In J. W. Guthrie (Ed.), *Encyclopedia of Education*. New York, USA: Macmillan Reference.

Schellens, T., Van Keer, H., & Valcke, M. (2004). Learning in asynchronous discussion groups: A multilevel approach to study the influence of student, group, and task characteristics. Manuscript submitted for publication.

Schellens, T., Van Keer, H., Valcke, M., & De Wever, B. (2005, June). The impact of role assignment as a scripting tool on knowledge construction in asynchronous discussion groups. *Proceedings of the Computer-supported Collaborative Learning conference 2005*, Taipei, Taiwan.

Schultz-Hardt, S., Jochims, M., & Frey, D. (2002). Productive conflict in group decision making: Genuine and contrived dissent as strategies to counteract biased information seeking. *Organizational Behavior & Human Decision Processes*, *88*, 563-586.

Sharan, S., & Shachar, H. (1988). *Language and learning in the cooperative classroom.* New York: Springer-Verlag.

Spada, H., Meier, A., Rummel, N., & Hauser, S. (2005, June). A new method to assess the quality of collaborative process in CSCL. *Proceedings of the Computer-supported Collaborative Learning conference 2005*, Taipei, Taiwan.

Stahl, G. (2005). Group cognition in computer-assisted collaborative learning. *Journal of Computer Assisted Learning*, *21*, 79-90.

Suthers, D. D. (2006). Technology affordances for intersubjective meaning making: A research agenda for CSCL. *International Journal of Computer-Supported Collaborative Learning*, *1*(3), 315-337.

Webb, N. (1984). Sex differences in interaction and achievement in cooperative small groups. *Journal of Educational Psychology*, *76*, 33-44.

Webb, N., & Kenderski, C. M. (1984). Student learning and interaction in small group and whole-class settings. In P. L. Peterson, L. C. Wilkinson, & M. Hallinan (Eds.), *The social context of instruction: Group organization and group processes* (pp. 153-170). Orlando: Academic.

Webster, M., Jr., & Driskell, J. (1983). Beauty as status. *American Journal of Sociology*, *89*, 140-165.

Zumbach, J., Schonemann, J., & Reimann, P. (2005, June). Analyzing and supporting collaboration in cooperative computer-mediated communication. Proceedings of the Computer-supported Collaborative Learning conference in Taipei, Taiwan.