

Chapter 1

A Complexity-Grounded Model for the Emergence of Convergence in CSCL Groups

Manu Kapur, John Voiklis, and Charles K. Kinzer

Abstract We advance a complexity-grounded, quantitative method for uncovering temporal patterns in CSCL discussions. We focus on convergence because understanding how complex group discussions converge presents a major challenge in CSCL research. From a complex systems perspective, convergence in group discussions is an emergent behavior arising from the transactional interactions between group members. Leveraging the concepts of *emergent simplicity* and *emergent complexity* (Bar-Yam 2003), a set of theoretically-sound yet simple rules was hypothesized: Interactions between group members were conceptualized as goal-seeking adaptations that either help the group move towards or away from its goal, or maintain its status quo. Operationalizing this movement as a Markov walk, we present quantitative and qualitative findings from a study of online problem-solving groups. Findings suggest high (or low) quality contributions have a greater positive (or negative) impact on convergence when they come earlier in a discussion than later. Significantly, convergence analysis was able to predict a group's performance based on what happened in the first 30–40% of its discussion. Findings and their implications for CSCL theory, methodology, and design are discussed.

1.1 Introduction

One of the major challenges facing collaborative problem-solving research is to understand the process of how groups achieve convergence in their discussions (Fischer and Mandl 2005). For example, without some level of convergent or shared assumptions and beliefs collaborators cannot define (perhaps not even recognize) the problem at hand nor select among the possible solutions, much less take action (Schelling 1960). A certain level of convergence—e.g., by what names collaborators refer to the objects of the problem—is required simply to carry

M. Kapur(✉)

National Institute of Education, 1 Nanyang Walk, Singapore 637616
e-mail: manu.kapur@nie.edu.sg

on a conversation (Brennan and Clark 1996). A deeper level of convergence—e.g., agreeing on the functional significance of the objects to which partners refer—is required to carry out shared intentions on those objects (Clark and Lucy 1975). Not surprisingly, Roschelle (1996) argued that convergence, as opposed to socio-cognitive conflict, is more significant in explaining why certain group discussions lead to more productive outcomes than others.

Although there has been considerable research towards understanding the cognitive and social mechanisms of convergence in collaborative learning environments (Clark and Brennan 1991; Fischer and Mandl 2005; Jeong and Chi 2007; Roschelle and Teasley 1995; Stahl 2005), the problem of convergence—understanding the complex process of how multiple actors, artifacts, and environments interact and evolve in space and time to converge on an outcome is critical—remains a perennial one (Barab et al. 2001). A substantial amount of literature attempts to understand group processes using qualitative analytical methods (e.g., interactional analysis, discourse analysis, conversation analysis), which provide insightful and meaningful micro-genetic accounts of the complex process of emergence of convergence in groups (e.g., Barron 2000, 2003; Teasley & Roschelle 1993; Stahl 2005). For the present purposes, however, our proposal speaks to quantitative approaches, typically involving quantitative content analysis (QCA) (Chi 1997) of interactional data; the use of QCA is pervasive in examining the nature of interaction and participation in CSCL research (Rourke Anderson 2004). However, quantitative measures and methods for conceptualizing the *temporal* evolution of collaborative problem-solving processes as well as the emergence of convergence remain lacking (Barab et al. 2001; Collazos et al. 2002; Hmelo-Silver et al. [this book](#); Reimann 2009).

Increasingly, a realization of the inherent complexity in the interactional dynamics of group members is giving way to a more emergent view of how groups function and perform (Arrow et al. 2000; Kapur et al. 2007; Kapur et al. 2008; Stahl 2005). However, the use of complex systems in the learning sciences is relatively sparse, but gaining momentum (see Jacobson and Wilensky 2006). A major thrust of such research is on the curricula, teaching, and learnability issues related to complex systems, and how they influence learning and transfer. However, complex systems also offer important theoretical conceptions and methodologies that can potentially expand the research tool-kit in the learning sciences (Jacobson and Wilensky 2006; Kapur and Jacobson 2009; Kapur et al. 2005, 2007). The work reported in this chapter leverages this potential to better understand how convergence emerges in group discussions.

From a complex systems' perspective, convergence in group discussions can be seen as a complex, emergent behavior arising from the local interactions between multiple actors, and mediated by tools and artifacts. Convergence is therefore a group-level property that cannot be reduced to any particular individual in the group. Yet, it emerges from and constrains the interactions between the very individuals it cannot be reduced to. To understand this emergence, we first discuss the concept of emergent behavior, particularly the distinction between *emergent simplicity* and *emergent complexity*; a distinction that is central to our proposal. Following that, we describe one way in which convergence in group discussions

can be conceptualized and modeled. We support our case empirically, through findings from a study of CSCL groups. We end by discussing the implications of our work for CSCL theory, methodology, and design.

1.1.1 Unpacking Emergent Behavior: Emergent Simplicity Versus Emergent Complexity

Central to the study of complex systems is how the complexity of a whole is related to the complexity of its parts (Bar-Yam 2003). The concept of emergent behavior—how macro-level behaviors emerge from micro-level interactions of individual agents—is of fundamental importance to understanding this relationship. At the same time, the concept of emergent behavior is rather paradoxical. On the one hand, it arises from the interactions between agents in a system, e.g., individuals in a collective. On the other hand, it constrains subsequent interactions between agents and in so doing, seems to have a life of its own independent of the local interactions (Kauffman 1995), and therefore, cannot be reduced to the very individual agents (or parts) of the system it emerged from (Lemke 2000). For example, social structures (e.g., norms, values, beliefs, lexicons, etc.) within social networks emerge from the local interactions between individual actors. At the same time, these structures shape and constrain the subsequent local interactions between individual actors, but they cannot be reduced to the very individual actors' behaviors and interactions they emerged from (Lemke 2000; Watts and Strogatz 1998). Therefore, it becomes fundamentally important to understand *how* macro-level behaviors emerge from and constrain micro-level interactions of individual agents.

Understanding the “how,” however, requires an understanding of two important principles in complexity. First, simple rules at the local level can generate complex emergent behavior at the collective level (Kauffman 1995; Epstein and Axtell 1996). For example, consider the brain as a collection of neurons. These neurons are complex themselves, but exhibit simple binary behavior in their synaptic interactions. This type of emergent behavior, when complexity at the individual-level results in simplicity at the collective-level, is called *emergent simplicity* (Bar-Yam 2003). Further, these simple (binary) synaptic interactions between neurons collectively give rise to complex brain “behaviors”—memory, cognition, etc.—that cannot be seen in the behavior of individual neurons. This type of emergent behavior, when simplicity at the individual-level results in complexity at the collective-level, is called *emergent complexity* (Bar-Yam 2003).

The distinction between emergent simplicity and complexity is critical, for it demonstrates the possibility that a change of scale (micro vs. macro level) can be accompanied with a change in the type (simplicity vs. complexity) of behavior (Kapur and Jacobson 2009). “Rules that govern behavior at one level of analysis (the individual) can cause qualitatively different behavior at higher levels (the group)” (Gureckis and Goldstone 2006, p. 1). We do not necessarily have to seek

complex explanations for complex behavior; complex collective behavior may very well be explained from the “bottom up” via simple, minimal information, e.g., utility function, decision rule, or heuristic, contained in local interactions (Kapur et al. 2006; Nowak 2004).

In this chapter, we use notions of emergent simplicity and complexity to conceptualize a group of individuals (agents) interacting with each other as a complex system. The group, as a complex system, consists of complex agents, i.e., just like the neurons, the individuals themselves are complex. Again, it is only intuitive to think that their behavior can be anything but complex and any attempt to model it via simple rules is futile. However, emergent simplicity suggests that this is not an ontological necessity. Their behavior may very well be modeled via simple rules. Further, emergent complexity suggests that doing so may reveal critical insights into the complexity of their behavior as a collective. It is this possibility that we explore and develop in the remainder of this chapter.

1.1.2 Purpose

We describe how convergence in group discussions can be examined as an emergent behavior arising from theoretically-sound yet *simple* teleological rules to model the collaborative, problem-solving interactions of its members. We support our model empirically, through findings from a study of groups solving problems in an online, synchronous, chat environment. Note that this study was part of a larger program of research on *productive failure* (for more details, see Kapur 2008, 2009, 2010; Kapur and Kinzer 2007, 2009). We first describe the context of the study in which the methodology was instantiated before illustrating the methodology.

1.2 Methodology

1.2.1 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 collaborate and solve either well- or ill-structured problem scenarios. The study was carried out in the school’s computer laboratory, 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.

1.2.2 Procedure

A well-structured (WS) and an ill-structured (IS) problem scenario were developed consistent with Jonassen's (2000) design theory typology for problems (see Kapur 2008 for the problem scenarios). Both problem scenarios dealt with a car accident scenario and targeted the same concepts from Newtonian Kinematics and Laws of Friction to solve them. 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. Pilot tests with groups of students from the previous cohort further informed the time allocation for the group work, which was set at 1.5 h.

The study was carried out in the school's computer laboratory. 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 1.5 h and solved their respective problems.

1.2.3 Hypothesizing Simple Rules

The concept of *emergent simplicity* was invoked to hypothesize a set of simple rules. Despite the complexity of the individual group members, the impact of their interactions was conceived to be governed by a set of simple rules. Group members were conceived as agents interacting with one another in a goal-directed manner toward solving a problem. Viewed *a posteriori*, these transactional interactions seemed to perform a telic function, i.e., they operated to reduce the difference between the current problematic state of the discussion and a goal state. Thus, local interactions between group members can be viewed as operators performing a

means-ends analysis in the problem space (Newell and Simon 1972). From this, a set of simple rules follows naturally. Each interaction has an impact that:

1. Moves the group *towards* a goal state, or
2. Moves the group *away* from a goal state, or
3. *Maintains* the status-quo (conceptualized as a “neutral impact”).

Then, convergence in group discussion was conceived as an *emergent complexity* arising from this simple-rule-based mechanism governing the impact of individual agent-based interactions.

1.2.4 Operationalizing Convergence

Concepts from the statistical theory of Markov walks were employed to operationalize the model for convergence (Ross 1996). Markov walks are commonly used to model a wide variety of complex phenomenon (Ross 1996). First, quantitative content analysis (QCA; Chi 1997), also commonly known as coding and counting, 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 vis-à-vis the hypothesized simple rule. Two trained doctoral students independently coded the interactions with an inter-rater reliability (*Krippendorff's alpha*) of .85. 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 1 s, -1 s, and 0 s.

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 convergence value was,

$$C = \frac{n_1 - n_{-1}}{n_1 + n_{-1}}.$$

For each of the 20 discussions, convergence values were calculated after each utterance in the discussion, resulting in a notional time series representing the evolution of convergence in group discussion.

1.3 Results

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

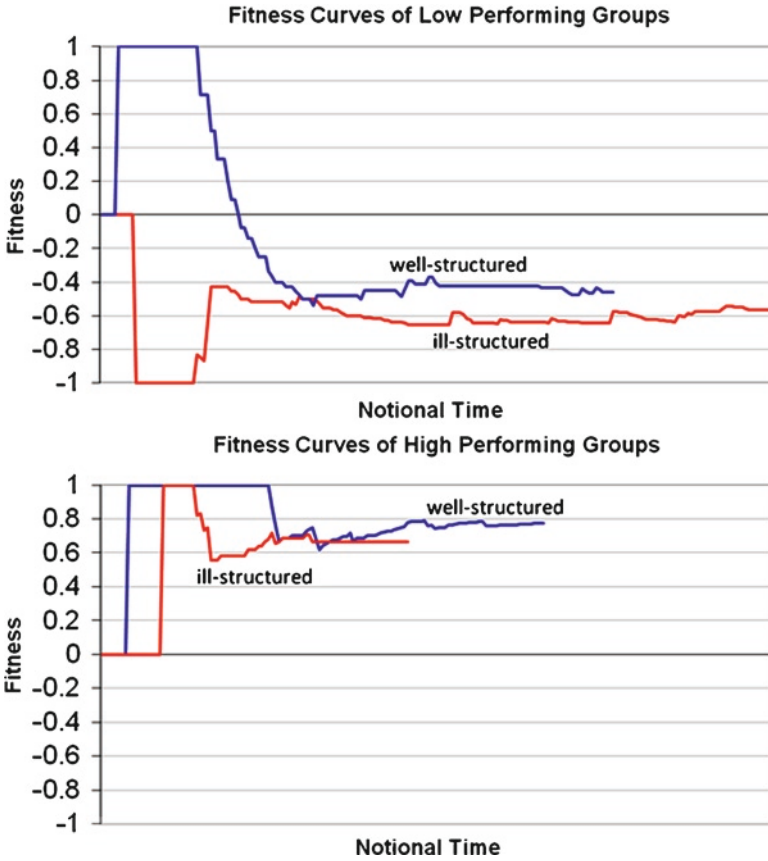


Fig 1.1 Fitness curves of high and low performing groups across problem types

in time. Figure 1.1 presents 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 (group performance is operationalized in the next section) across the well- and ill-structured problem types.

1.3.1 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 one. However, the data

suggests that, in reality, some level of divergence of ideas may in fact be a good thing (Kapur 2008, 2009; Schultz-Hardt et al. 2002), 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 Fig. 1.1, 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 well-structured group showed a sharp fall after initially moving in the correct direction (indicated by high fitness initially). The ill-structured 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 with the low-performing groups, 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. What is most interesting is that this descriptive examination of fitness curves provides a view of paths to a solution that are lost in analysis systems that 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 consider only single points in time (often only the solution state) are not consistent with what this study's data suggest about problem-solving 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 one. However, the denominator in the ratio always increases, regardless of the contribution being positive or negative. Therefore, when a positive (or negative) contribution comes earlier in the discussion, its impact on convergence is greater because a unit increment (or 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 following 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.

1.3.2 Relationship between Convergence and Group Performance

The purpose of the simulation was to determine if the level of convergence in group discussion provided an indication of the eventual group performance. Group performance was operationalized as the quality of group solution, independently rated by two doctoral students on a nine-point rating scale (Table 1.1) with an inter-rater reliability (*Krippendorff's alpha*) of .95.

Table 1.1 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

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

The discussions of all 20 groups were each segmented into ten equal parts. 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 (well- or ill-structured) 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 Fig. 1.2).

C1 through C10 denote the ten 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 Fig. 1.2), 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. This insight bears important implications for scaffolding group discussions to achieve optimal outcomes. For example, if one's primary aim is to maximize 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 higher fitness plateaus; the higher the fitness plateau, the better the group performance, on average. This insight is in and of itself a significant finding, but since participation in high-performing groups is consistently (and not surprisingly) a strong predictor of subsequent individual learning gains (e.g., see Barron 2003; Cohen et al. 2002), we believe it makes it all the more significant, for it demonstrates strong connections to group and individual learning.

To delve deeper into what makes convergence a powerful measure, micro-analytical interactional analysis sheds more light. We present a brief analysis of the following excerpt containing an exchange between group members S1 and S2.

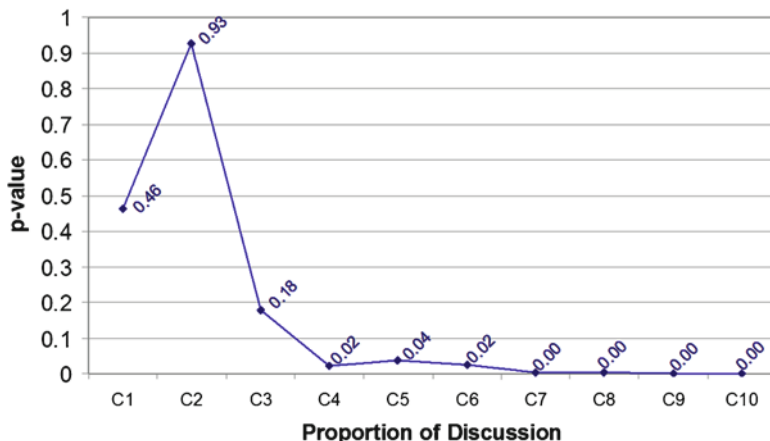


Fig. 1.2 Simulation of the significance of convergence in predicting group performance

<i>S1</i> >are we going to apply frictional retardation for the reaction time also?	-1
<i>S2</i> >no, because reaction time is the time between watching the boy and applying the brakes so in this time [the] car must be accelerating	1, 1
<i>S1</i> >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
<i>S2</i> >but maximum speed is at the instant when he applied the brake	1
<i>S1</i> >but earlier you said that the car will accelerate after perceiving the boy	-1
<i>S2</i> >I said so because his foot must be on accelerator during reaction time	1
<i>S1</i> >Now I understand... please proceed to write your answer	1, 1

Recall that the problem involved a car-accident scenario (see Kapur 2008, for the problem 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 the driver’s reaction time. Being a misconception, it was rated as having a negative impact (-1). S2 evaluates S1’s question and says ‘no,’ attempting to correct the misconception. Hence, its positive (+1) impact rating. In the same utterance, S2 elaborates why frictional retardation should not to be applied, further positively impacting the group’s progress. 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. Note that had S2 wrongly evaluated and agreed to S1’s misconception, the impact ratings would have been negative, which, without any further correction, would have led the group to diverge from a correct understanding of that very aspect of the problem.

This analysis, albeit brief, shows that impact ratings are meaningful only in relation to preceding utterances (Bransford and Nitsch 1978) and take into account the

sequence and temporality of collaborative interactions (Kapur et al. 2008). 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. As a result, despite solving different types of problems (well- or ill-structured), group performance depended mainly upon the convergence of their discussions. Because convergence takes into account both the number as well as the temporal order of the units of analyses, it utilizes a greater amount of information present in the data. This makes convergence a more powerful measure, both conceptually and statistically, than existing predictors that do not fully utilize the information present in interactional data. If this is the case, then the following hypothesis should hold: *convergence is a more powerful predictor of group performance than existing, commonly-used interactional predictors.*

1.3.3 Comparing Convergence with Other Commonly-Used Interactional Predictors

Many studies of collaborative problem solving, including this one, use QCA to operationalize measures for problem-solving processes. These measures typically result in data about the frequency or relative frequency of positive indicators (e.g., higher-order thinking, questioning, reflecting, etc.), or negative indicators (e.g., errors, misconceptions, lack of cooperation, non-task, etc.), or a combination that adjusts for both the positive and negative indicators (e.g., the difference between the frequencies of high- and low-quality contributions in a discussion). In this study, we operationalized three measures to represent typical measures:

1. Frequency (n_1 : recall that this is the number of interaction units in a discussion with impact=1),
2. Relative Frequency, $\left(\frac{n_1}{n_1 + n_0 + n_{-1}}\right)$, and
3. Position, ($n_1 - n_{-1}$).

Convergence, $\left(\frac{n_1 - n_{-1}}{n_1 + n_{-1}}\right)$, formed the fourth measure.

Multiple linear regression was used to simultaneously compare the significance of the four measures in predicting group performance, controlling for problem type in each case. The overall model was significant, $F=6.391$, $p=.003$. Results in Table 1.2 suggest that, of the four predictors of group performance, convergence was the only one significant, $t=2.650$, $p=.019$, thereby supporting our hypothesis. In other words, consistent with our hypothesis, convergence seems to be a more powerful predictor of problem-solving performance when compared to existing, commonly-used predictors.

Table 1.2 Regression parameter estimates of the interactional variables and problem type

	B	SE	Beta	t	p
(Constant)	-1.778	1.693		-1.051	.311
Frequency	.006	.021	.239	.268	.793
Relative frequency	1.541	2.792	.116	.552	.590
Position	-.012	.024	-.446	-.513	.616
Convergence	5.338	2.014	.839	2.650	.019
Problem type	-.050	.544	-.021	-.092	.928

1.4 Discussion

In this chapter, we have described a complexity-grounded model for convergence. We argued that convergence in group discussions can be seen as a complex, emergent behavior arising from the local interactions between multiple actors, and mediated by tools and artifacts. That is, convergence is a group-level property that cannot be reduced to any particular individual in the group. Yet, it emerges from and constrains the interactions between the very individuals it cannot be reduced to. A complexity-grounded model allowed us to model a complex, group-level emergent behavior such as convergence using simple interactional rules between group members. More specifically, we used the concepts of emergent simplicity and emergent complexity to hypothesize a set of theoretically-sound yet simple rules to model the problem-solving interactions between group members, and then examined the resulting emergent complexity: Convergence in their discussion.

Despite the intentional simplicity of our model, it revealed novel insights into the process of collaboration. The first insight concerned the differential impact of contributions in a group discussion—high (or low) quality contributions have a greater positive (or negative) impact on the eventual outcome when they come earlier than later in a discussion. A corollary of this finding was that eventual group performance could be predicted based on what happens in the first 30–40% of a discussion because group discussions tended to settle into fitness plateaus fairly quickly. Finally, convergence was shown to be a more powerful predictor of group performance than some existing, commonly-used measures. These insights are significant, especially since participation in high-performing groups is a strong predictor of subsequent individual learning gains (e.g., see Barron 2003; Cohen et al. 2002). In other words, this conceptualization and analysis of convergence demonstrates strong connections to both group performance and individual learning.

1.4.1 Implications for Scaffolding

If, as our work suggests, group performance is highly sensitive to early exchanges in a discussion, then this insight bears important implications for scaffolding synchronous, small-group, CSCL discussions to achieve optimal outcomes. For example, if

one primarily aims to maximize group performance, these early sensitivities suggest 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 higher fitness plateaus; better group performance is predicated by high fitness plateaus. This is also consistent with the notion of fading, that is, having scaffolded the early exchange, the scaffolds can be faded (Wood et al. 1976). 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 problem categorization processes often occur early in problem-solving discussions and can shape all subsequent processes (Kapur and Kinzer 2007; Kapur et al. 2005; Voiklis 2008). Such an approach stands in contrast with the 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.

1.4.2 Implications for Methodology: The Temporal Homogeneity Assumption

Sensitivity to early exchange also underscores the role of temporality, and consequently, the need for analytical methods to take temporal information into account. According to Reimann (2009), “Temporality does not only come into play in quantitative terms (e.g., durations, rates of change), but order matters: Because human learning is inherently cumulative, the sequence in which experiences are encountered affects how one learns and what one learns” (p. 1). Therefore, understanding (1) how processes evolve in time, and (2) how variation in this evolution explains learning outcomes, ranks among the more important challenges facing CSCL research (Akhraas and Self 2000; Hmelo-Silver et al. *this book*; Reimann 2009).

To derive methodological implications, let us first consider a prototypical case of coding and counting in CSCL. Typically, one or more coding/rating schemes are applied to the interaction data, resulting 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 a certain category of interaction? It could be that interactions coded in that category were spread throughout the discussion, or perhaps they were clustered together in a coherent phase during the discussion. Therefore, interactions 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. By aggregating category counts over time, one implicitly makes the assumption of *temporal homogeneity* (Kapur et al. 2008). 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 (Voiklis et al. 2006).

It follows, then, that we need methods and measures that take temporality into account. These methods and measures can potentially allow us to uncover patterns in time and reveal novel insights (e.g., sensitivity to early exchange) that may otherwise not be possible. Consequently, these methods and measures can play an instrumental role in the building and testing of a process-oriented theory of problem solving and learning (Reimann 2009).

1.4.3 Implications for Theorizing CSCL Groups as Complex Systems

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 (Arrow et al. 2000; Bar-Yam 2003); 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 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 (Hmelo-Silver et al. *this book*). 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 (Voiklis et al. 2006).

At a more conceptual level, the idea that one can derive meaningful insights into a complex interactional process via a simple rule-based mechanism, while compelling, may also be unsettling and counter-intuitive. Hence, a fair amount of intuitive resistance to the idea is to be expected. For instance, it is reasonable to argue that the extreme complexity of group interaction—an interweaving of syntactic, semantic, and

pragmatic structures and meanings operating at multiple levels—make it a different form of emergence altogether and, therefore, insights into complex interactional processes cannot be gained by using simple-rule-governed methods. However, a careful consideration of this argument reveals an underlying ontological assumption that complex behavior cannot possibly be explained by simple mechanisms. Saying it another way, some may argue that *only* complex mechanisms (e.g., linguistic mechanisms) can explain complex behavior (e.g., convergence in group discussion). Of course, this is a possibility, but, notions of emergent simplicity and emergent complexity suggest that this is not the “only” possibility (Bar-Yam 2003), especially given our knowledge of the laws of self-organization and complexity (Kauffman 1995).

It is noteworthy that emergent complexity is also integral to the theory of dynamical minimalism (Nowak 2004) used to explain complex psychological and social phenomena. Dynamical minimalist theory attempts to reconcile the scientific principle of parsimony—that simple explanations are preferable to complex ones in explaining a phenomenon—with the arguable loss in depth of understanding of that phenomenon because of parsimony. Using the principle of parsimony, the theory seeks the simplest mechanisms and the fewest variables to explain a complex phenomenon. It argues that this need not sacrifice depth in understanding because simple rules and mechanisms that repetitively and dynamically interact with each other can produce complex behavior: the very definition of emergent complexity. Thus, parsimony and complexity are not irreconcilable, leading one to question the assumption that complex phenomena necessarily require complex explanations (Kapur and Jacobson 2009).

Therefore, the conceptual and methodological implication from this study is not that complex group behavior ought to be studied using simple-rule-based mechanisms, but that exploring the possibility of modeling complex group behavior using simple rule-based mechanisms is a promising and meaningful endeavor. Leveraging this possibility, we demonstrated one way in which simple-rule-based mechanisms can be used to model convergence in group discussion, in turn revealing novel insights into the collaborative process. The proposed measures of convergence and fitness curves were intentionally conceived and designed to be generic and, therefore, may be potentially applicable to other problem-solving situations as well. Thus, they also provide a platform for the development of more sophisticated measures and techniques in the future.

1.4.4 Some Caveats and Limitations

New methods and measures always raise more questions than answers, and ours is not an exception. What is more important is that repeated application and modification over multiple data sets is needed before strong and valid inferences can be made (Rourke and Anderson 2004). At this stage, therefore, our findings remain technically bound by the context of this study; it is much too early to attempt any

generalization. There are also several issues that need to be highlighted when considering the use of the proposed methodology:

Issues of coding: Clearly, drawing reliable and valid inferences based on the new measures minimally requires that the coding scheme be reliable and valid. To this end, a conscious, critical decision was our choice of the content domain: we chose Newtonian kinematics because it is a relatively well-structured domain. This domain structure clarified the task of differentiating those contributions to the problem-solving discussion that moved the group closer to a solution from those that did not and, thus, minimized the effort of coding the impact of interactional units (1, -1, and 0). For a more complex domain (e.g., ethical dilemmas) where the impact is not as easy to assess, our method may not be as reliable, or perhaps not even applicable.

Model simplicity: It can be argued that the proposed model is a very simple one, and could be seen as a limitation. The decision to keep the model simple was intentional; we chose to keep the number of codes to a minimum, i.e., just -1, 0, or 1. We reasoned that if we could reveal novel insights by using the simplest model, then one could always “complexify” the model subsequently. For example, one could easily build on this model to code impact on a five point scale from -2 to 2 so as to discriminate contributions that make a greater positive or negative impact than others. At the same time, the model also collapses many dimensions (such as social, affective, cognitive, meta-cognitive, and so on) into one dimension of impact. Collapsing dimensions into a simple model allows for the easy and direct interpretation of results, but this gain in interpretability comes with the cost of an overly reductive model. Once again, we wanted to demonstrate that even with a simple model, one could potentially gain insights, and having done so, one could always embark on a building a more complex model. For example, it might be useful to model the co-evolution of the various dimensions, investigate the co-evolving fitness trajectories, and develop deeper understandings of the phenomenon.

Corroborating interpretations: Our model is essentially a quantitative model. In reducing complex interactions into impact ratings, it is necessarily reductive. In interpreting findings from such analyses, it is important to use the method as part of a mixed-method analytical commitment. If not, it may be hard to differentiate results that are merely a statistical or mathematical artifact from the ones that are substantively and theoretically meaningful.

1.4.5 Future Directions

Going forward, we see the need for developing new temporal measures. We want to focus particularly on those that can be easily implemented from a temporal sequence of codes that QCA of group discussions normally results in. In particular, we argue for two candidates:

1. **Lag-Sequential Analysis (LSA):** LSA treats each interactional unit (as defined in a study) as an observation; a coded sequence of these observations forming the

interactional sequence of a group discussion (e.g., Erkens et al. 2003). LSA detects the various non-random aspects of interactional sequences to reveal how certain types of interactions follow others more often than what one would expect by chance (Wampold 1992). By examining the transition probabilities between interactions, LSA identifies statistically significant transitions from one type of interactional activity to another (Bakeman and Gottman 1997; Wampold 1992). As a result, the collaborative process can be examined as an evolving, multi-state network, thereby allowing us to reveal temporal patterns that may otherwise remain hidden (Kauffman 1995). For example, Kapur (2008) coded collaborative problem solving interactions into process categories of problem analysis, problem critique, criteria development, solution development, and solution evaluation, thereby reducing each group discussion into a temporal string of process category events. Using LSA, the analysis revealed significant temporal patterns that the typical coding and counting method could not reveal, that is, how some process categories were more likely to follow or be followed by other process categories significantly above chance level. More importantly, LSA demonstrated how variation in temporal patterns—sequences of process categories—was significantly related to variation in group performance.

2. **Hidden Markov Models (HMMs):** HMMs (Rabiner 1989) offer analysis at a relatively coarser grain size than LSA by detecting the broader interaction phases that a discussion goes through. For example, Soller and colleagues (2002) used HMMs to analyze and assess temporal patterns in on-line knowledge sharing conversations over time. Their HMM model could determine the effectiveness of knowledge sharing phases with 93% accuracy, that is, 43% above what one would expect by chance. They argued that understanding these temporal phases that provide an insight into the dynamics of how groups share, assimilate, and build knowledge together is important in building a process theory of facilitating to increase the effectiveness of the group interactions. Conceiving a group discussion as a temporal sequence of phases, one can use several methods to isolate evolutionary phases, including measures of genetic entropy (Adami et al. 2000), intensity of mutation rates (Burtsev 2003) or, in the case of problem interactions, the classification of coherent phases of interaction. With the phases identified, one can calculate and predict the probabilities between phases using HMMs (Rabiner 1989; for an example, see Holmes 1997). As a result, one may begin to understand when and why phase transitions as well as stable phases emerge; more importantly, one may begin to understand how the configuration of one phase may influence the likelihood of moving to any other phase. Whether one can control or temper these phases, or whether such control or temperance would prove a wise practice remains an open question which, even if only partially answered, will be a breakthrough in characterizing and modeling the problem solving process.

It is worth reiterating that these methods should not be used in isolation, but as part of a larger, multi-method, multiple grain size analytical program. At each grain size, findings should potentially inform and be informed by findings from analysis at other grain sizes—an analytical approach that is commensurable with the

multiple levels (individual, interactional, group) at which the phenomenon unfolds. Only then can these methods and measures play an instrumental role in the building and testing of a process-oriented theory of problem solving and learning (Hmelo-Silver et al. [this book](#); Reimann 2009; Reimann et al. [this book](#)).

1.5 Conclusion

In this chapter, we have advanced a complexity-grounded, quantitative method for uncovering temporal patterns in interactional data from CSCL discussions. In particular, we have described how convergence in group discussions can be examined as an emergent behavior arising from theoretically-sound yet simple teleological rules to model the collaborative, problem-solving interactions of its members. We were able to design a relatively simple model to reveal a preliminary yet compelling insight into the nature and dynamics of problem-solving CSCL groups. That is, convergence in group discussion, and consequently group performance, is highly sensitive to early exchanges in the discussion. More importantly, in taking these essential steps toward understanding of how temporality affects CSCL group processes and performance, we call for further efforts within this line of inquiry.

Acknowledgments The work reported herein was funded by a Spencer Dissertation Research Training Grant from Teachers College, Columbia University to the first author. This chapter reports work that has, in parts, been presented at the International Conference of the Learning Sciences in 2006, and the Computer-Supported Collaborative Learning Conference in 2007. Our special thanks go to the students and teachers who participated in this project. We also thank June Lee and Lee Huey Woon for their help with editing and formatting.

References

- Adami, C., Ofria, C., & Collier, T. C. (2000). Evolution of biological complexity. *Proceedings of the National Academy of Sciences*, 97, 4463–4468.
- Akhras, F. N., & Self, J. A. (2000). Modeling the process, not the product, of learning. In S. P. Lajoie (Ed.), *Computers as cognitive tools* (No more walls, Vol. 2, pp. 3–28). Mahwah: Erlbaum.
- Arrow, H., McGrath, J. E., & Berdahl, J. L. (2000). *Small groups as complex systems: Formation, coordination, development, and adaptation*. Thousand Oaks: Sage.
- Bakeman, R., & Gottman, J. M. (1997). *Observing interaction: An introduction to sequential analysis*. New York: Cambridge University Press.
- Barab, S. A., Hay, K. E., & Yamagata-Lynch, L. C. (2001). Constructing networks of action-relevant episodes: An in-situ research methodology. *Journal of the Learning Sciences*, 10(1&2), 63–112.
- Barron, B. (2000). Achieving coordination in collaborative problem-solving groups. *Journal of the Learning Sciences*, 9, 403–436.
- Barron, B. (2003). When smart groups fail. *Journal of the Learning Sciences*, 12(3), 307–359.
- Bar-Yam, Y. (2003). *Dynamics of complex systems*. New York: Perseus.

- Bransford, J. D., & Nitsch, K. E. (1978). Coming to understand things we could not previously understand. In J. F. Kavanaugh & W. Strange (Eds.), *Speech and language in the laboratory, school, and clinic* (pp. 267–307). Harvard: MIT Press.
- Brennan, S. E., & Clark, H. H. (1996). Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22(6), 1482–1493.
- Burtsev, M. S. (2003). Measuring the dynamics of artificial evolution. In W. Banzhaf, T. Christaller, P. Dittrich, J. T. Kim, & J. Ziegler (Eds.), *Advances in artificial life. Proceedings of the 7th European conference on artificial life* (pp. 580–587). Berlin: Springer.
- Chi, M. T. H. (1997). Quantifying qualitative analyses of verbal data: A practical guide. *Journal of the Learning Sciences*, 6(3), 271–315.
- Clark, H., & Brennan, S. (1991). Grounding in communication. In L. B. Resnick, J. Levine, & S. D. Teasley (Eds.), *Perspectives on socially shared cognition* (pp. 127–149). Washington, DC: APA.
- Clark, H. H., & Lucy, P. (1975). Understanding what is meant from what is said: A study in conversationally conveyed requests. *Journal of Verbal Learning and Verbal Behavior*, 14(1), 56–72.
- 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.
- Collazos, C., Guerrero, L., Pino, J., & Ochoa, S. (2002). *Evaluating collaborative learning processes. Proceedings of the 8th international workshop on groupware (CRIWG'2002)* (pp. 203–221). Berlin: Springer.
- Epstein, J. M., & Axtell, R. (1996). *Growing artificial societies: Social science from the bottom up*. Washington, DC/Harvard: Brookings Institution Press/MIT Press.
- Erkens, G., Kanselaar, G., Prangmsma, M., & Jaspers, J. (2003). Computer support for collaborative and argumentative writing. In E. De Corte, L. Verschaffel, N. Entwistle, & J. van Merriënboer (Eds.), *Powerful learning environments: Unravelling basic components and dimensions* (pp. 157–176). Amsterdam: Pergamon, Elsevier Science.
- Fischer, F., & Mandl, H. (2005). Knowledge convergence in computer-supported collaborative learning: The role of external representation tools. *Journal of the Learning Sciences*, 14(3), 405–441.
- Gureckis, T. M., & Goldstone, R. L. (2006). Thinking in groups. *Pragmatics and Cognition*, 14(2), 293–311.
- Hmelo-Silver, C. E., Jordan, R., Liu, L., & Chernobilsky, E. (this book). Representational tools for understanding complex computer-supported collaborative learning environments. In S. Puntambekar, G. Erkens, & C. Hmelo-Silver (Eds.), *Analyzing interactions in CSCL: Methods, approaches and issues* (pp. 83–106) Springer.
- Holmes, M. E. (1997). Optimal matching analysis of negotiation phase sequences in simulated and authentic hostage negotiations. *Communication Reports*, 10, 1–9.
- Jacobson, M. J., & Wilensky, U. (2006). Complex systems in education: Scientific and educational importance and implications for the learning sciences. *Journal of the Learning Sciences*, 15(1), 11–34.
- Jeong, H., & Chi, M. T. H. (2007). Knowledge convergence during collaborative learning. *Instructional Science*, 35, 287–315.
- Jonassen, D. H. (2000). Towards a design theory of problem solving. *Educational Technology Research and Development*, 48(4), 63–85.
- Kapur, M. (2008). Productive failure. *Cognition and Instruction*, 26(3), 379–424.
- Kapur, M. (2009). Productive failure in mathematical problem solving. *Instructional Science*. doi:10.1007/s11251-009-9093-x.
- Kapur, M. (2010). A further study of productive failure in mathematical problem solving: Unpacking the design components. *Instructional Science*. DOI: 10.1007/s11251-010-9144-3.
- Kapur, M., Hung, D., Jacobson, M., Voiklis, J., Kinzer, C., & Chen, D.-T. (2007). Emergence of learning in computer-supported, large-scale collective dynamics: A research agenda. In C. A. Clark, G. Erkens, & S. Puntambekar (Eds.), *Proceedings of the international conference of computer-supported collaborative learning* (pp. 323–332). Mahwah: Erlbaum.

- Kapur, M., & Jacobson, M. J. (2009). Learning as an emergent phenomenon: Methodological implications. Paper presented at the annual meeting of the American educational research association. San Diego.
- Kapur, M., & Kinzer, C. (2007). The effect of problem type on interactional activity, inequity, and group performance in a synchronous computer-supported collaborative environment. *Educational Technology Research and Development*, 55(5), 439–459.
- Kapur, M., & Kinzer, C. (2009). Productive failure in CSCL groups. *International Journal of Computer-Supported Collaborative Learning*, 4(1), 21–46.
- Kapur, M., Voiklis, J., & Kinzer, C. (2005). Problem solving as a complex, evolutionary activity: A methodological framework for analyzing problem-solving processes in a computer-supported collaborative environment. In *Proceedings the computer supported collaborative learning (CSCL) conference*. Mahwah: Erlbaum.
- Kapur, M., Voiklis, J., & Kinzer, C. (2008). Sensitivities to early exchange in synchronous computer-supported collaborative learning (CSCL) groups. *Computers & Education*, 51, 54–66.
- 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: Erlbaum.
- Kauffman, S. (1995). *At home in the universe: The search for the laws of self-organization and complexity*. New York: Oxford University Press.
- Lemke, J. L. (2000). Across the scales of time: Artifacts, activities, and meanings in ecosocial systems. *Mind, Culture and Activity*, 7(4), 273–290.
- Newell, A., & Simon, H. (1972). *Human problem solving*. Englewood Cliffs: Prentice Hall.
- Nowak, A. (2004). Dynamical minimalism: Why less is more in psychology. *Personality and Social Psychology Review*, 8(2), 183–192.
- Rabiner, L. R. (1989). A tutorial on hidden markov models and selected applications in speech recognition. *Proceedings of the IEEE*, 77(2), 257–286.
- Reimann, P. (2009). Time is precious: Variable- and event-centred approaches to process analysis in CSCL research. *International Journal of Computer-Supported Collaborative Learning*, 4(3), 239–257.
- Reimann, P., Yacef, K., & Kay, J. (this book). Analyzing collaborative interactions with data mining methods for the benefit of learning. In S. Puntambekar, G. Erkens, & C. Hmelo-Silver (Eds.), *Analyzing interactions in CSCL: Methods, approaches and issues* (pp. 161–185). Springer.
- Roschelle, J. (1996). Learning by collaborating: Convergent conceptual change. In T. Koschmann (Ed.), *CSCL: Theory and practice of an emerging paradigm* (pp. 209–248). Mahwah: Erlbaum.
- Roschelle, J., & Teasley, S. D. (1995). The construction of shared knowledge in collaborative problem solving. In C. E. O'Malley (Ed.), *Computer-supported collaborative learning* (pp. 69–197). Berlin: Springer.
- Ross, S. M. (1996). *Stochastic processes*. New York: Wiley.
- Rourke, L., & Anderson, T. (2004). Validity in quantitative content analysis. *Educational Technology Research and Development*, 52(1), 5–18.
- Schelling, T. C. (1960). *The strategy of conflict*. Cambridge: Harvard University Press.
- 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 and Human Decision Processes*, 88, 563–586.
- Soller, A., Wiebe, J., & Lesgold, A. (2002). A machine learning approach to assessing knowledge sharing during collaborative learning activities. In G. Stahl (Ed.), *Proceedings of computer support for collaborative learning* (pp. 128–137). Hillsdale: Erlbaum.
- 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.

- Teasley, S. D., & Roschelle, J. (1993). Constructing a joint problem space: The computer as a tool for sharing knowledge. In S. P. Lajoie & S. D. Derry (Eds.), *Computers as Cognitive Tools* (pp. 229–258). Hillsdale, NJ: Erlbaum.
- Voiklis, J. (2008). A thing is what we say it is: Referential communication and indirect category learning. PhD thesis, Columbia University, New York.
- Voiklis, J., Kapur, M., Kinzer, C., & Black, J. (2006). An emergentist account of collective cognition in collaborative problem solving. In R. Sun (Ed.), *Proceedings of the 28th annual conference of the cognitive science society* (pp. 858–863). Mahwah: Erlbaum.
- Wampold, B. E. (1992). The intensive examination of social interaction. In T. R. Kratochwill & J. R. Levin (Eds.), *Single-case research design and analysis: New directions for psychology and education* (pp. 93–131). Hillsdale: Erlbaum.
- Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of “small-world” networks. *Nature*, 393, 440–442.
- Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 17, 89–100.