Tracing Ideas and Participation in an Asynchronous Online Discussion across Individual and Group Levels over Time

Alyssa Friend Wise, Ying-Ting Hsiao, Farshid Marbouti, Yuting Zhao
Simon Fraser University, 250-13450 102 Avenue, Surrey BC, Canada, V3T 0A3
Email: afw3@sfu.ca, yha73@sfu.ca, fmarbout@sfu.ca, yza174@sfu.ca

Abstract: This paper advances a methodology to support a coordinated multi-level analysis of individual actions and group dynamics in asynchronous online discussions. The approach uses log-file data to examine group and individual participation patterns, and argumentation coding of post contents to probe developmental trajectories of individual and collective understandings. Importantly, these traces of ideas and behaviors are coordinated within and across levels. To illustrate the method, the paper presents an analysis of five undergraduate students taking part in a five-day online discussion to address a business challenge. Findings provided insight into the ways in which phenomena at the group and individual levels interrelated and drove each other, specifically the complex interdependency between group and individual willingness to engage in debate, and how one dominant individual’s unfounded presumption of consensus led to early abandonment (rather than conscious rebuttal) of ideas.

Introduction
Asynchronous text-based discussion forums are often employed to facilitate online collaboration among students. Temporal flexibility gives students time to reflect on their ideas and those of others, thus theoretically affording opportunities for the group to collectively construct knowledge and for individuals to develop their personal understanding (Stahl, 2005). However, in practice, studies have found that learners often exhibit low cognitive engagement (Hew et al., 2010) with shallow and unfocused argumentation (Nussbaum, 2005). The problems of superficial discussion and weak argumentation can be considered on both the individual level (students don’t engage with or challenge each other’s ideas) and that of the group (teams engage in consensus building without considering multiple ideas first). These levels are inherently interconnected as the group’s process emerges from the accumulation of individual actions while also situating and constraining them (Suthers & Teplov, 2011). Similarly, efforts to support productive dialogue in online discussions may be targeted at one of these levels, but will necessarily have interdependent effects on both. For example, requiring a discussion group to come to a consensus (group-level intervention) will affect the ways individuals contribute to the conversation; conversely giving students specific responsibilities in a discussion (individual-level intervention) will impact their group dynamics. For this reason, understanding how a pedagogical intervention affects participation and argumentation processes in online discussions requires a coordinated examination of individual- and group-level phenomena as they evolve over time.

Attention to the multiple levels at which collaborative learning can be studied is growing (Stahl & Hesse, 2008) as is the use of statistical techniques that account for both individual and group influences on collaboration (Cress, 2008). However little work has attempted to examine how phenomena at these levels mutually depend on and affect each other. In addition, recent research in collaborative learning has emphasized the importance of studying the flow of student interactions over time (Reimann, 2009). This paper advances a methodology to support a coordinated multi-level temporal analysis of individual actions and group dynamics in online discussions. To illustrate the method, we present an analysis of a group of five students taking part in a five-day online discussion to reach consensus on how to address an organizational business challenge.

A Coordinated Multi-level Microanalytic Approach
To understand how conversation flows in an asynchronous online discussion we track trajectories of participation and ideas across group and individual levels and examine how they interrelate over time. The detailed traces of student activity created in online discussions present a rich source of data for such examination. Specifically, log-file data of actions (e.g. how students make posts and access those of others) facilitate analysis of group and individual participation patterns, while the content of posts supports examination of collective and individual understandings. While previous approaches have studied these data sources in aggregate (e.g. Palmer et al., 2008), we have found that unpacking data to provide an on-the-ground view of particular students and groups provides insight not available from collectivized data. This work draws on our previous research developing different analytic techniques to study how groups and individuals participate and develop understanding over time (Wise & Chiu, 2011; Wise et al., 2012). The important addition here is the attempt to coordinate analyses across levels. Such coordination is challenging for several reasons. First, while a group’s participation is spread over hours and days, individuals experience participation in discrete episodes lasting minutes; thus differing timescales must be considered (Lemke, 2000). Second, the artifact of a group’s
complete discussion is not how the living conversation appeared to individual participants at any given point in time. Finally, there is the difficulty of inferring relationships between phenomena that exist at different ontological levels. We address these challenges through three sequential analytical steps.

**Log-file data analysis**

Log-file data is extracted by participant. Three types of actions are coded: “views” (opening others’ posts), “posts” (creating a post), and “reviews” (revisiting one’s own posts). Durations are calculated by subtracting time between subsequent actions. Views are further categorized as scans or reads using a cut-off maximum reading speed of 6.5 words per second (Hewitt et al., 2007). Adjustments are made for unlogged system exits and actions are grouped into sessions (a series of consecutive actions). To trace the group’s timeline of participation, we visualize the log-file data to show patterns of group participation over threads and time (Wise & Padmanabhan, 2009; see Fig. 1); an interpretive narrative is then created. Separately, we create a narrative reconstructing the experience for each individual action-by-action using a dynamic discussion map to provide the context of how a discussion appeared at the time a behavior was taken. The combination of these data sources supports a meaningful reconstruction of individuals’ behaviors in the forum (Wise et al., 2012).

**Post content analysis**

Our examination of post contents draws on work from the field of argumentation. A coding scheme adapted from Clark and Sampson (2008) is used to evaluate when students propose, support and rebut ideas using the post as the unit of analysis (multiple codes can be applied to each post). The scheme also captures if students analyze the problem and if they attempt to synthesize the group’s position. With this approach we can look at how the group negotiates ideas to move to a common conclusion, how individuals’ ideas change over time, and how these relate to the ways in which students position themselves with respect to the group’s dialogue. Other schemes and models of knowledge construction could certainly be used within our overall methodology as well. Once posts are coded, we create a collective visual map of positions on each idea over time to trace the group’s idea trajectory (see Fig. 2a). To examine the evolution of individuals’ ideas, a second map for each student is drawn (see Fig. 2b). Both diagrams are interpreted to create initial idea development narratives.

**Coordination of data sources within and across levels**

While log-file and post content data are initially analyzed separately, the data streams are highly interrelated. For example, an individual’s new idea might be driven by their actions (e.g. reading a certain post) and highly contested ideas might trigger a collective burst of activity. Thus our first coordination step is to integrate the analyses of actions and ideas at each level (individual and group) into coherent narratives. We then use the narratives at each level to inform each other: contextualizing the participation and ideas of each individual within the group’s trajectory, and tracing how patterns of group activity emerged from actions of particular individuals. Thus for a group of N students, the result is N+1 accounts of a discussion, one from the perspective of each student and one for the group as a whole. While the narratives at each level necessarily refer to each other, they are distinct in the primacy given to the perspective of the group or a specific individual.

**Sample Analysis**

We present a sample analysis using our multi-level microanalytic methodology with data from a group of five students (Arlene, Ben, Celia, David and Evan) solving a collaborative task that was seeded with divergent opinions but required consensus. Each student’s trajectory of participation and ideas was traced along an individual timeline, as was the collective process of participation and argumentation for the group as a whole. These parallel narratives were then compared and coordinated to examine the ways in which processes and phenomena at the individual and group levels interrelated and drove each other.

**Learning Context and Participants**

Data was taken from a larger study comparing the impact of different task-types on online discussion in a blended undergraduate business course. Students were randomly assigned to small groups, each responsible for discussing one business challenge online (Friday to Tuesday) and leading a subsequent tutorial discussion. The challenge in the week studied was communication. A junior executive was told that her response to criticism of her presentation by senior executives was too forceful. Two possible solutions were given for discussion: either email the senior executives to explain why she spoke up or accept the feedback and adjust her future behavior; additional solutions could also be proposed. Students were asked to come to consensus on a solution with a rationale using theories from class. Posting guidelines and criteria (worth 5% of course grade) were provided.

**Process and Findings**

Log-file data from the chosen discussion was extracted and processed as described above. Contents of all posts were coded by two researchers; agreement was 89% or higher for all categories except problem analysis (73%).
Group and individual timelines of ideas and participation were visualized in diagrams. Due to space limitations in a short paper, in this presentation we focus on the participation of three students (Arlene, Ben, Celia) over two key days as an illustration of the insight that the multi-level microanalytic approach can provide.

**The Group**

The group’s discussion focused primarily on three solutions, two provided in the task (accept the feedback [Accept] and email the executives [Email]) and a variation on the latter of these (talk to the executives in person [Talk]). The participation peak for the discussion (greatest number of posting and reading actions) occurred on Sunday (see Fig. 1); however argumentation in the discussion was greatest one day prior in a back-and-forth sequence between Ben and Arlene debating the merits of the three solutions. All three other students also logged-in on this day and read the debate posts, but none joined in the debate. This appears to have been a pivotal moment for the group’s discussion; the choice of other students to not join the debate while it was “hot” rendered what could have been a central locus of argumentation peripheral in the discussion. On Sunday many posts were made in support of the three solutions, but none rebutted the opposing alternatives. That evening, Arlene made a post presuming the group’s consensus was Talk. This was a second pivotal moment for the group as all subsequent comments unquestioningly accepted this as the group’s solution and focused on how to implement it or how to present it in the class tutorial; none questioned its status as the solution. Thus the group’s argumentation was effectively completed on Sunday evening without clear rebuttal of all unchosen alternatives.

**Arlene**

On Friday, Arlene initiated the first thread analyzing the problem and listing three possible solutions (Accept, Email, Talk) without indicating a preference. On Saturday, she logged in the forum and spent just one minute to read and respond to a post by Ben supporting Email. Her position in this post seems to have developed in reaction to Ben’s stance; she rebutted his argument by pointing out that the executives could easily misinterpret an email, thus she advocated Talk. That evening, Arlene read and replied back to a new counterargument Ben had made by restating the possibility of miscommunication via Email and her preference for Talk. However, she also suggested that maybe no action was necessary (Accept). On Sunday afternoon, Arlene returned to the
discussion, quickly skimmed a few posts and then posted in a thread analyzing the problem in light of upward/downward communication theory. An hour later, she logged-in again and immediately started a new thread that attempted to connect her preferred solution (Talk) to her problem analysis. She posted “earlier we agreed there are 3 main problems… [but] I just realized that in our solution, we actually only address problem #2… The face-to-face meeting with the senior exec might clear up [this] misunderstanding.” This post assumed that the rest of the group agreed with her about the problems identified (“we agreed”) though two of these had only been mentioned by her, and about the solution of Talk (“our solution”) though a group consensus had not been explicitly discussed and both support and dissent for Talk had been expressed. Her reasons for presuming consensus are unclear; it may be that she took Ben’s abandonment of the debate as tacit agreement for Talk, or that since she spent very little time reading others’ ideas she was unaware that other solutions still had support. In this post she also shifted away from her earlier support of Accept as a possible solution without explanation.

Ben
Ben first contributed to the discussion on Saturday, reading all the previously made posts and taking a stand for Email for the employee to “clear-up any misunderstandings… and show that she is capable of what she is doing.” That evening he spent almost a half hour reading a rebuttal responded by Arlene and composing his own counterclaim in support of Email stating that Talk was unpractical based on the theory of power distance in organizations. He also rebutted another peer’s post suggesting to Accept the feedback since this would leave a bad impression of the employee. Ben logged in again later that night and read Arlene’s subsequent rebuttal of his defense; however, he did not continue the debate with her. On Sunday, Ben was active in reading others’ posts, but not in making his own. He made only one post, and in it did not advocate for or against any solutions. This might be a result of his minority position in the discussion, a reaction to the general lack of dissent, or his realization (revealed in a later post accepting Talk) that the group should agree on a solution. This final possibility suggests Ben interpreted the need for consensus as a reason not to debate different options.

Celia
Celia made her first post on Saturday evening, after spending considerable time that day and the one before reading the posts of others. Though the debate between Ben and Arlene was in progress, she did not join in, instead starting a new thread analyzing the problem. On Sunday morning, Celia again spent time reading others’ posts and then made another post emphasizing that the group should identify the problem before trying to solve it. After, she responded to the arguments in the Ben-Arlene debate thread by explaining her preference for Accept, but also indicating support for Arlene’s idea of Talk. While she did not support Email, she also did not refute it; instead she expressed understanding about why Ben might take that position. Thus Celia’s response to the debate dynamic was diplomatic, acknowledging the value of each of the ideas involved. Celia did not log in again till Monday night, when she spent the better part of an hour catching up the discussion before writing a short reply accepting the group’s consensus of Talk.

Discussion and Conclusions
The coordination of multiple levels of analysis helps shed light on specific ways in which individual actions influenced the group’s process and how the group collectively set the stage for individual actions. Close to the start of discussion one student took a stand for a solution, prompting another to rebut his idea and initiate a debate dynamic at the group level. However the reluctance of others to join-in, perhaps because they valued reaching consensus over persuading others of their opinions (Nussbaum, 2005), created a collective lack of dissent. Left in the minority, the first student abandoned his argument. The following day the group shifted to initiating many new threads. This created a disjointed discussion, making it possible for one dominant student’s unjustified presumption of consensus to go uncontested. As a result, the group stopped considering other solutions, a deviation from dialogical argumentation (Clark & Sampson, 2008). While one student seemed to equally influence the group and be influenced by it, another was insistent in driving the group, and the third seemed content to follow the group’s will. This highlights differences in the reciprocity of influence between the group and individuals. Several effects of timing also seemed to mediate: first, comments of individuals who were active early set the group’s focus; second, the richest argumentation occurred in a short time. This raises questions about the value of compressed / distributed participation and the temporal weighting of voice.

This paper presented a methodology for coordinated multi-level analysis to examine interrelationships between individual actions and group dynamics in online discussions. The methodology is still in development; testing and revision with additional data is needed. While the approach appears fruitful for generating insight into relationships between individuals’ trajectories of behavior and understanding, and groups’ patterns of dialogue and idea-development, its fine-grained, interpretive and time-intensive nature raises important questions about intersubjectivity, generalizability and scalability. Here, we present our current thinking around these issues. First, with respect to intersubjectivity, our initial participation narratives and diagrams are based on log-file data and constructed objectively as neutral, straight-forward reports of the temporal sequence of actions.
in which an individual engaged or in which posts were contributed by the group. Content analysis of posts involves researcher subjectivity; we address this through the well-established practices of using a clear and theoretically-grounded coding scheme, choosing an unambiguous unit of analysis, and reporting inter-rater reliability (de Wever et al., 2006). The major interpretive steps occur when the initial idea and participation analyses for an individual or the group are integrated into coherent narratives and when these interpretive narratives are used to contextualize and inform each other. Making inferences about how different kinds of entities (ideas and behaviors) affect each other and relationships between phenomena that exist at different ontological levels inherently engages the researcher as meaning-maker and thus necessarily involves attention to intersubjectivity. For this reason, the interpretive narratives are repeatedly negotiated by multiple researchers throughout their creation, review and revision.

Nonetheless, as with any fine-grained analysis that focuses on a small number of participants, the goal is not to make statistical generalizations or suggest that what was found in one particular case will likely be found in another. Instead, the objective is theoretical generalization (Seale, 1999); that is to use the deep examination of a particular occurrence to develop theoretical understandings that can present a useful lens for interpreting other situations and suggest implications for educational practice. For example, this sample analysis highlighted the importance of how and why ideas stop being considered, the notion of reciprocity between group and individual levels; and the importance of temporal sequence and pace. Finally there is the question of scalability. The multi-level microanalytic approach involves multiple steps of technical data processing, content analysis, and interpretation. The time to conduct the complete analysis of this one discussion by five students over five days is estimated at 150 hours. Thus, use on a large scale is prohibitively expensive in terms of man-hours. Instead, this methodology might be most effectively used in combination with existing temporally-sensitive large-scale methods (e.g. Chiu & Khoo, 2005) which can identify important moments or phases in collaboration worthy of the investment in in-depth multi-level microanalysis.

References

Acknowledgments
This work was supported by the Social Sciences and Humanities Council of Canada.