Analyzing the Organization of Collaborative Math Problem-solving in Online Chats using Statistics and Conversation Analysis

Alan Zemel[1], Fatos Xhafa[2] and Gerry Stahl[3]

[1]The Math Forum, Drexel University, 3210 Cherry Street,
Philadelphia, PA USA
arz26@drexel.edu

[2]The Open University of Catalonia, Barcelona, Spain
fxhafa@uoc.edu

[3]College of Information Science and Technology,
Philadelphia, PA USA
gerry.stahl@cis.drexel.edu

Abstract. In this paper we describe how a statistical test on a hypothesis regarding collaborative math problem solving using online chats showed an unexpected result, whose understanding required the use of qualitative methods. The phenomenon behind the result is identified using Conversation Analysis. This paper demonstrates the importance of using qualitative methods to describe the perspective of participants as a way of interpreting statistical results, revising hypotheses and developing alternative coding schemes and procedures. The combined approach of quantitative and qualitative methods is applied on real data coming from Virtual Math Teams research project (Drexel University) and is identifying issues not addressed so far in the analysis of online collaborative group activity.

1 Introduction

The analysis of the use of groupware is particularly problematic. Most methods of human-computer interaction were developed for single-user systems and are not applicable to computer mediation of group interaction. A common approach to analyzing the use of groupware is to compare statistical measures of usage across conditions or cases. However, this can be criticized for not investigating and taking into account qualitative differences that may be crucial to understanding the quantitative differences [1]. While there is a widespread feeling that fields like CSCL and CSCW need to take a multidisciplinary approach incorporating a variety of analytic methods, it is difficult to see how quantitative and qualitative approaches built on fundamentally incompatible theoretical foundations can work together. This paper reports a case in which a quantitative discovery led to qualitative analysis that explained the significance of the quantitative results and suggested modifications of the quantitative approach.

In the Virtual Math Teams (VMT, [2]) project at Drexel University, we investigate online problem-solving chat interactions from a variety of analytical and methodological perspectives. On the one hand, a coding scheme has been developed and applied to logs of online chats among actors participating in math problem solving. This provides a basis for a quantitative analysis of the chat logs. On the other hand, conversation analytic methods have been applied to these chat logs as a way of describing the procedures participants use to make sense of their ongoing activity.

Conversation analysis (CA) and statistical analysis (SA) are uneasy partners in the analytic enterprise. These two orientations to analysis derive from very different perspectives on the role of the analyst and the kinds of assumptions that can be made with respect to the data and its interpretation. In statistical analysis, hypotheses are put forward and tested. Coding schemes are devised which are designed to facilitate the testing of these hypotheses and statistical methods are applied to coded data. In this approach, it is the analyst's perspective that is privileged. The analyst:

- proposes the hypotheses,
- produces the coding scheme to capture the relevant data from an experiment designed specifically to allow for testing of the hypothesis, and
- assesses and interprets the statistical results [3].

Statistical analysis of data gathered from online collaborative learning experiments plays a central role in many CSCL studies [4], [5], [6], [7]. A whole range of statistical methods, from descriptive statistics to multilevel and other sophisticated methods have been used to analyze the underlying features (variables) of the collaborative activity that takes place in a small group.

Conversation analysis, on the other hand, is an analytical methodology that attempts to describe the actions of participants in terms of the relevances demonstrated by participants in and as their interaction [8], [9]. This methodology privileges the perspective of the participants over the analyst's perspective [10]. Actions are seen as situated within a stream of ongoing action and are sequentially organized. Furthermore, conversation analysts presume that actors design and 'customize' their action for the particular circumstances in which they are accomplished.

The differences between SA and CA are consequential. For statistical analysts, validity and reliability are significant concerns. These are not concerns for conversation analysts. Conversation analysts are concerned with providing adequate descriptions of the sense-making procedures used by participants as they interact. Where statistical analysts would discover what might be 'present' as frequently observed regularities in interactions, conversation analysts are concerned with how specific actions are made relevant by prior actions and how a current action make relevant subsequent actions over the course of a particular sequence of actions. For conversation analysts, it is sufficient that the participants in a particular interaction treat their ongoing actions as sensible. The conversation analyst's task is to describe these sequences of actions as sense-making procedures.

While these two types of analysis, statistical and conversational, may seem incompatible, it turns out there are circumstances in which they can be mutually informative [11]. In this paper, we describe a situation in which a puzzling statistical result was made intelligible by conversation analytic investigation. This is a novel approach to analyze the organization of the interaction in collaborative math problem-solving activities in online chats. Indeed, existing approaches in the literature treat quantitative and qualitative methods separately. Our results show the strength of using a combined approach. Specifically, by using a quantitative approach, we detected an unexpected result in a hypothesis test. This made further investigation necessary. The qualitative method enabled us to identify the phenomenon that produced the unexpected result in the hypothesis test.

2 Data Collection

The Virtual Math Teams (VMT, [2]) project at Drexel University investigates small group collaborative learning in mathematics. In this project an experiment is being conducted, called *powwow*, which extends The Math Forum's (mathforum.org) "*Problem of the Week (PoW)*" service. Groups of 3 to 5 students in grades 6 to 11 collaborate online synchronously to solve math problems that require reflection and discussion. AOL's Instant Messenger software is used to conduct the experiment in which each group is assigned to a chat room. Each session lasts about one to one and a half hour. The *powwow* sessions are recorded as chat logs (transcripts) with the handle name (the participant who made the posting), timestamp of the posting, and the content posted.

2.1 Coding Scheme

Both quantitative and qualitative approaches are employed in the VMT project to analyze the transcripts in order to understand the interaction that takes place during collaboration within this particular setting. A coding scheme has been developed in the VMT project to quantitatively analyze the sequential organization of interactions recorded in a chat log. The unit of analysis is defined as one posting that is produced by a participant at a certain point of time and displayed as a single posting in the transcript.

The coding scheme includes nine distinct dimensions, each of which is designed to capture a certain type of information from a different perspective. They can be grouped into two main categories: one is to capture the content of the session whereas another is to keep track of the threading of the discussion, that is, how the postings are linked together. Among the content-based dimensions, conversation and problem solving are two of the most important ones which code the conversational and problem solving content of the postings. Related to these two dimensions are the Conversation Thread and the Problem Solving Thread, which provide the linking between postings, and thus introduce the relational structure of the data. The conversation thread also links fragmented sentences that span multiple postings. The problem solving thread aims to capture the relationship between postings that relate to each other by means of their mathematical content or problem solving moves (see Figure 1).

Line #	Handle	Statement	Time	Conversation Thread	Conversation	Problem Solving Thread	Problem Solving
45	AVR	Okay, I think we should start with the formula for the area of a triangle	8:21:46		Offer		Strategy
46	SUP	ok	8:22:17	45	Follow	45	
47	AVR	A = 1/2bh	8:22:28		Offer	45	Perform
48	AVR	I believe	8:22:31	47	Extension	47	
49	PIN	yes	8:22:35	47	Setup	47	
50	PIN	i concue	8:22:37	49	Agree	49	Check
51	PIN	concur*	8:22:39	50	Repair Type		
52	AVR	then find the area of each triangle	8:22:42		Offer	45	Strategy
53	AVR	oh, wait	8:22:54		Regulation		
54	SUP	the base and heigth are 9 and 12 right?	8:23:03		Request		Offer
55	AVR	no	8:23:11	54	Setup	54	
56	SUP	o	8:23:16		No Code		
57	AVR	that's two separate triangles	8:23:16	55	Critique	55	Reflect
58	SUP	ooo	8:23:19	55	Setup	55	
59	SUP	ok	8:23:20	58	Response	58	

Figure 1: A coded excerpt from Pow2a.

Each dimension has a number of subcategories. The coding is done manually by 3 trained coders independently after strict training assuring a satisfactory reliability. Regarding the statistical approach, this paper considers 4 dimensions only; namely the conversation, problem solving, social reference, math move and system support dimensions.

2.2 Data Collection

The sample used in this study consisted of six powwows that were chosen from a larger set of powwows with the aim at conducting a first data analysis. The criteria for choosing the sample is based on one of the characteristics of the powwow experiment, namely, for some powwows the math problem was announced in advance while for some others the math problem was announced just at the time of starting the chat session[1]. Thus, the sample of six powwows is made up of three powwows in which the math problem was announced at the beginning of the session, whereas in the rest the problem was posted on the Math Forum's web site in advance[2] (see Table 1). It should be noted that for the math problem being announced in advance doesn't necessarily mean that the participants of the chat already solved the problem in advance.

---

[1] We will refer to this as "*known – not known*" criterion.

[2] We will refer to the first group as "*NO group*" and to the second as "*YES group*."

Table 1: Description of the coded chat logs.

PoW-wow Session #	Facilitator	Members	Number of Utterances	PoW Name	Announced Before?
1	MUR	PIN, GOR, REA, MCP	334	Finding CE	No
2a	GER	FIR, PIN, SUP, OFF	724	Equilateral Triangle Areas	No
2b	MUR	MCP, AH3, REA	204	Equilateral Triangle Areas	No
9	POW	EEF, AME, AZN, LIF, FIR	715	Making triangles	Yes
10	MFP	AME, FIR, MCP	582	The perimeter of an octagon	Yes
18	MFP	AME,KOH,KIL,ROB	488	A tangent square and circle	Yes

1. Statistical Analysis

First Level: Statistical Analysis Based on Main Dimensions

Our first objective was to test whether there is any significant effect of the “known – not known” criterion on the sample of the powwows. To this end, we started by computing[3], through descriptive statistics, the distribution of frequencies in different dimensions (*Conversation*, *Social Reference*, *Problem Solving*, *Math Move* and *System Support*) for the six Powwows and used Means and ANOVA[4] to test the existence of significance difference due to the “known – not known” criterion. The study showed that there was no such effect, at a usual confidence level of 95% (in fact, significance in differences, that is significant pairs, were not noticed even at 90% confidence level). The fact that there is no clear effect of the criteria “known – not known” allows us to conclude that the classification of the sample of Powwows into groups according to “known – not known” criterion is not relevant. We could also observe this by computing the Boxplot representation of the variables under study (see Figure 2).

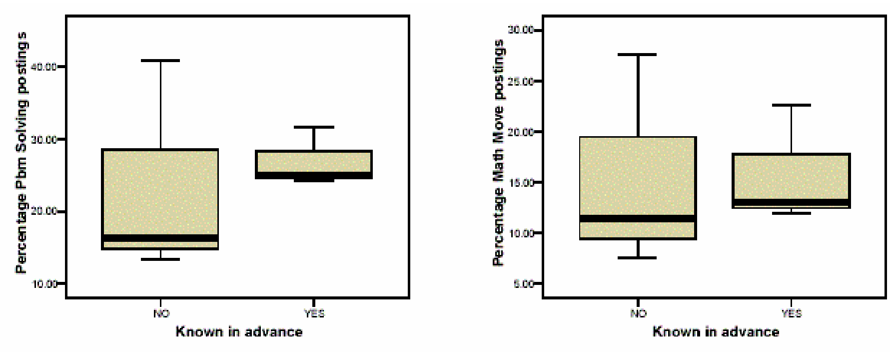

Figure 2: Boxplot representation of Problem Solving and Math Move dimensions.

---

[3] The statistical computations are done in SPSS 12.0.
[4] Note that the different dimensions are independent of each other

Given the above finding, we refined the statistical analysis by looking at the correlation between vectors of values of the six powwows (grouping "known – not known" was now maintained just for visual effect). By computing similarities between the powwows we could see which powwows are similar to each other and which are different from each other. We computed thus the correlations (Pearson correlations) through proximity matrix shown in Table 2.

Table 2. Pearson Correlation of Vector Values of Six Powwows

Proximity Matrix

	Correlation between Vectors of Values					
	1:NO	2:NO	3:NO	4:YES	5:YES	6:YES
1:NO	1.000	0.756	-0.452	0.567	0.108	-0.197
2:NO	0.756	1.000	-0.219	0.912	0.603	0.067
3:NO	-0.452	-0.219	1.000	0.202	0.620	0.956
4:YES	0.567	0.912	0.202	1.000	0.867	0.470
5:YES	0.108	0.603	0.620	0.867	1.000	0.791
6:YES	-0.197	0.067	0.956	0.470	0.791	1.000

This is a similarity matrix

From Table 2 we observe the following:

a) Pow2b (3:NO in the table) is *negatively* correlated to the powwows of the NO group (pow1 and pow2a) and *positively* correlated to the powwows of the other group (pow9, pow10, pow18). Moreover, significant correlation of pow2b with pow10 (5:YES) and pow18 (6:YES) is observed and not significant correlation with the pow9 (4:YES).

b) There is a significant positive correlation of the pow9 with pow1 and pow2a of the NO group. In pair wise terms, pow9 is more correlated to the powwows of the NO group than to the powwows of the YES group (its group).

c) There are some pairs of powwows positively and strongly correlated, namely (powwow2a, pow9) and (pow2b, pow18) which suggest taking a closer study of the possible common features of these powwows.

The previous observations on the correlations between powwows from different groups not only supports the claim that there is no significant effect of the "known – not known" criterion but also shed light on the reason why these two groups are not really separated. Indeed, the negative correlation of the pow2b with the powwows of the NO group shows that its place is not in the NO group. Even more, its positive correlation with the powwows of the YES group indicates that this powwow is better grouped with the powwows of the YES group.

In our next step, we decided to exclude the *System Support* dimension from the analysis; indeed, this dimension is less relevant in the context of the interaction analysis and could have thus introduced some noise in the analysis. We run again the statistical computations by re-computing the correlations through proximity matrix as shown in Table 3.

Table 3. Pearson Correlation of Vector Values of Six Powwows (system support excluded)

Proximity Matrix

	Correlation between Vectors of Values					
	1:NO	2:NO	3:NO	4:YES	5:YES	6:YES
1:NO	1.000	0.999	-0.427	0.868	0.376	-0.145
2:NO	0.999	1.000	-0.396	0.884	0.407	-0.112
3:NO	-0.427	-0.396	1.000	0.080	0.678	0.957
4:YES	0.868	0.884	0.080	1.000	0.787	0.366
5:YES	0.376	0.407	0.678	0.787	1.000	0.862
6:YES	-0.145	-0.112	0.957	0.366	0.862	1.000

This is a similarity matrix

By excluding the System Support dimension, we observe a clear effect on the correlations, namely:

a) On the one hand, an increased negative correlation of the pow2b (3:NO) with the powwows of its group (pow1 and pow2a, 1:NO and 2:NO, respectively) is now observed. Notice also that the correlation between pow1 and pow2a is almost perfect correlation. One the other hand an increased positive correlation of the pow2b (3:NO) with the powwows of the other group (pow9, pow10, pow18) is observed. Interestingly, pow2b is now less correlated to pow9 (4:YES in the table).

b) An increased positive correlation of Pow9 with the powwows of the NO group (pow1 and pow2a) is now observed. Moreover, we observe a decrease in its correlation with pow10 and pow18.

c) Finally, pow18 is now negatively correlated to both pow1 and pow2a.

We repeated the above computations by standardizing the variable values by z-score.

Table 4. Proximity Matrix

	Correlation between Vectors of Values					
	1:NO	2:YES	3:YES	4:NO	5:YES	6:YES
1:NO	1.000	.987	-.999	.869	-.921	-.993
2:NO	.987	1.000	-.977	.778	-.845	-.999
3:YES	-.999	-.977	1.000	-.894	.939	.986
4:NO	.869	.778	-.894	1.000	-.993	-.808
5:YES	-.921	-.845	.939	-.993	1.000	.870
6:YES	-.993	-.999	.986	-.808	.870	1.000

This is a similarity matrix

According the statistical computations indicated above, the powwows show the following two clusters:

Cluster 1: (pow1, pow2a, pow9)
Cluster 2: (pow2b, pow10, pow18)

By re-computing[5] the Boxplot representation of this new clustering we could observe the significant separation between variables under study for the two groups (see Figure 3).

In other words, we expected the chat logs to be clustered based on the idea that in some chats, participants had access to the problem prior to their participation in the chat, while in other chats, participants had no access to the problem. However, the statistical analysis demonstrated that the clustering of chats was organized according to some other basis. At this point, we determined to conduct a qualitative approach to identify the reasons for this alternative organization of the online chats.

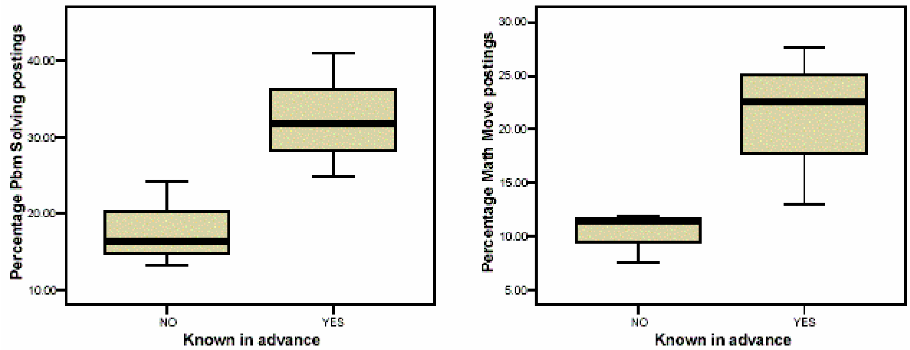

Figure 3: Boxplot representation of Problem Solving and Math Move dimensions

[5] Compare to Figure 2.

2. Participation Frameworks and the Organization of Online Interaction

To discover possible reasons for the failure of our initial hypothesis, we reexamined the chats using Conversation Analysis (CA). With this approach, we examined logs of the online chats to identify participants' perspectives on their own actions with an eye to describing their actions as sense-making procedures. The work of conversation analysis involves close inspection of interactional data. In conventional face-to-face interaction, this involves inspecting video and audio recordings of interaction. When it comes to online chats, logs of the chats, which display the text postings of participants and the time stamp associated with each posting, are the data that are inspected.

"Conversation analysis studies the order/organization/orderliness of social action, particularly those social actions that are located in everyday interaction, in discursive practices, in the sayings/tellings/doings of members of society" ([9], p.2). The object of inquiry in conversation analysis is not exclusively conversation per se, but rather talk and social interaction. Thus, as Ten Have describes, "CA's interest is with the local production of [social] order and with 'members' methods' for doing so" ([8], p.19).

Using the methods of CA, we began to notice that the organization of social order in these chats could be differentiated according to the way that participants oriented to the production of problem solutions. In particular, we noticed that, in some circumstances, participants reported on *work they had already completed*, whether it was work done prior to the chat or work done offline and without the participation of others in the production of that work during a chat. This organization of participation we have termed **expository participation**. On the other hand, we noticed that there were circumstances in which participants engaged each other in the *conjoint discovery and production* of both the problem and possible solutions. This organization of participation we termed **exploratory participation**.

Expository participation in the chats we examined involved one actor producing a report as an extended narrative of an activity performed by that actor. Such reporting is designed to project recipient participation in terms of the production of assessments of the report or the reported work. Recipients of that report have not participated in the work being reported. The report is designed and presented either as an already achieved understanding of the problem in terms of a candidate solution or as steps anyone with appropriate understanding of the problem might take to produce a solution. One version of expository participation is where one actor first announces that a solution has been achieved and then, upon prompting from recipients, proceeds to tell recipients what the solution is and how he or she produced the candidate solution. For example, an actor might report something like: "I've got the answer" which calls upon recipients to solicit the result. Announcing a result makes it relevant for recipients to ask for an explanation. Explanations might be offered in ways that describe the production of the solution as having been already achieved by the actor reporting the result, as in, "First I did … and then I computed … which equals …" Another way to produce an explanation involves the circumstance where an actor describes how a competent person would go about solving the problem, as in "First you do … then you compute … which equals …” In this regard, these approaches to the exposition of a problem’s solution is much like the telling of a story (see e.g. [12]). This is illustrated in the chat excerpt below:

Fig. 2. Example of Expository Chat (Powwow 2b)

24	AH3	I think I have the solution!
25	REA	what
26	MCP	I guess 15
27	REA	k
28	MCP	I think it's like the Pythagorean idea, applying to triangles.
29	AH3	sqrt(5^2 + 7^2) = sqrt(74)
30	MCP	Yes, 30-60-90 is needed fact
31	AH3	The solution is sqrt(74)
32	REA	how
33	MCP	7?
34	AH3	Go to...
35	AH3	http://mathforum.org/dr.math/faq/form ulas/faq.triangle.html
36	AH3	Under scalene triangle, the formula for the area of any triangle is...
37	AH3	K = a^2 * sin(B) * sin(C)/[2 sin (A)]
38	AH3	Why is that smiley their
39	AH3	K = a^2 * sin(B)* sin(C)/[2 sin (A)]
40	AH3	Where a = an edgelength of an isosceles triangle

An expository report is a way that an actor constitutes a problem as solvable. This is, in fact, a position we support because there is evidence in the transcripts that actors themselves orient to these reports in just this way. For example, the actor producing the report treats the problem as having already been solved and thereby constitutes a participation framework in which that he or she acts in the manner of an instructor, explaining what is already known by the instructor to an audience that presumably does not yet know. Constituting such a participation framework is a delicate business in the conduct of these chats. To do so, actors often draw upon the resources of news reporting by indicating they have something newsworthy to report, i.e., the solution to the problem. The actor reporting the solution designs his or her report in a way that allows the recipients of the report to “discover” in the report how the problem can be seen as solvable and solved.[6]

Exploratory participation, on the other hand, involves participation in which actors interact so as to constitute, in and as their chat, an understanding of a problem in terms of the conjoint production of possible organizations of mathematical activity from which a solution could be achieved. In such circumstances, actors use the resources afforded them by their interaction to constitute the math problem and their understanding of that problem as an emergent sequence of possible and/or achieved math activities designed to produce what may come to be subsequently recognizable

[6] This is similar to the way Livingston finds mathematicians doing proofs [13].

and treated as a solution to the problem. If expository participation is a form of "news" reporting, then the distinguishing feature of exploratory participation is that the actors themselves are constituting the "news" as their ongoing interaction rather than reporting it and receiving the report. This is shown below:

Fig. 3. Example of Exploratory Chat (Powwow 2b)

119	MCP	What's this extra saying? Like, if both of the smaller triasngles are sitting on their bases, the base of one is 5 and the base of the other is 7? Is that the interpretation?
120	REA	I guess it is 10
121	MCP	If they're oriented with corresp angles in corresp locations?
122	REA	or should I say 9.8
123	REA	what do you think
124	REA	i used the proportions
125	MCP	Oh, I guess this is where that 7 from AH3's answer came from, way back there. I didn't know where that came from.
126	MCP	I still need to make sure I know what the wording is saying. Am I interpreting the q right?

Actors engaged in exploratory work do not have a solution in hand as they do when they are engaged in expository work. Instead, they work to discover ways to produce such a solution by 1) allocating participation among actors in the chat and 2) by constituting and drawing on resources for producing a solution that are distributed among participants and which are made available by actors' participation in the chat. Like expository participation, the work of exploratory participation also constitutes the problem in terms of its solution, but with exploratory participation the solution is not yet known to participants. Exploratory interactions involve putting forward proposals for consideration and assessment, negotiating ways of formulating the problem in terms of different solution strategies, quick exchanges among multiple participants rather than extended postings, etc. Thus the work of exploration involves something developing alternative understandings of the problem in terms of the development and assessment of alternative possible solutions.

It is important to note that expository and exploratory work may be done during the same chat. Furthermore, expository participation requires that the expositor did the work of producing a solution "offline," i.e. without the participation of other actors in the chat. One of the affordances of chat is that "offline" activities are possible even as a chat is occurring because participants only have access to the messages they post. An actor's work with a pencil and a pad of paper beside his or her computer is not available to others unless and until it is posted in the chat system.

By examining the Powwow chats, we were able to see that there were considerable differences in the way participation was organized. Despite the fact that actors in Powwow 2b had not seen the problem in advance of their chat, they did their work "offline" during the chat and displayed an expository organization of participation in common with Powwows 10 and 18. Despite the fact that the actors in Powwow 9 had access to the problem in advance of the chat, they displayed an exploratory organization of participation in common with Powwows 1 and 2a. Thus, using CA, we were able to identify the same relation among the powwows showed by the statistical analysis and, moreover, explain the phenomenon in terms of the organization of participation in the chats.

One important question we considered was whether or not the coding scheme that had been used to identify these puzzling clusters initially could have been used to identify these different organizations of participation. We decided that it would not have been possible. The reasons for this decision are as follows. The existing coding scheme treated each post as the primary unit of analysis. Codes applied to individual chat postings but could not be used to characterize larger sequences of postings. This made it impossible to analytically identify the organization of participation which is understood as the relation among groupings of posted chat messages. While an alternative approach to coding might have made such an analysis possible, the work involved in developing such a coding scheme was formidable. Furthermore, it pointed to the logical problem of consistency that coding schemes are often designed in ways that lend themselves to find things for which there are codes. If we want to understand how participants organize their participation, if we want to understand a sequence of actions from participants' perspectives, then coding schemes need to capture these perspectives rather than the perspectives and interests of the researcher.

3. Hypothesis Testing and Discovery

As this work has shown, analytically understanding social interaction can be a tricky business. The conduct of inquiry into social interaction has traditionally utilized theories and analytical methodologies that allowed the analyst to test hypotheses against a collection of coded data [14]. By proposing hypotheses and testing them against coded data derived from "real world" phenomena, analysts are presumed to be able to check the validity of their theories about social interaction. On occasion, anomalies appear. Unexpected results are either dismissed as "outliers" or other methods of analysis are deployed to provide some explanation.

Many of the problems associated with statistical analysis of social phenomena derive from the coding schemes that are used. The procedures for producing codes and for applying them to interactional data are sometimes problematic. One problem is that the sense-making procedures analysts use to produce and apply codes are not independent of the sense-making procedures participants in the observed social interaction use to make sense of their ongoing activity. While this can be seen as a failure of coding schemes, it can also be viewed as a resource for doing initial investigations of social interactional phenomena which are then supplemented by close inspection of the sense-making procedures actors use. This is the perspective and approach we have taken in the VMT project.

Using the existing coding scheme as applied to six chat logs, we explored the hypothesis that we would expect to see a difference in the way that problem-solving chats were organized if participants did or did not have the opportunity to inspect the problem in advance of the chat. We hypothesized a difference could be detected and used statistical techniques to describe ways that the chats were grouped together. What we found was counter to the hypothesis. Rather than dismiss the results or question the value of the coding scheme, we opted to treat the unexpected result as an indicator of phenomena that required further investigation and closer analysis of the way that participants organized their activities in these chats.

4. Conclusions and further work

In this research, we were able to exploit the mutually informing features of quantitative and qualitative analysis. This has allowed us to discover a far more nuanced explanation for the observed grouping of chats. However, in order to determine whether our qualitative results provide an adequate explanation across multiple cases, we need to re-specify a coding scheme that derives from the perspective of the participants (for further discussion, see [14], [11]). According to [11] practitioners of CA have often made informal distributional claims with respect to observed interactional phenomena. However, certain questions about the 'typicality' or distribution of certain features of interactions of a particular type can only be assessed quantitatively. In such cases, questions arise as to the appropriate way to code data such that the requirements of valid statistical and quantitative analysis can be met without violating the requirements of preserving the sequential organization of, participants' perspectives on and relevances with respect to emergent, unfolding action sequences.

Based on this research, we have begun to explore a 'top-down' approach to coding, based on the ways that interactants organize themselves and their interaction into recognizable activities. This approach uses CA methods to identify closings and openings of action sequences by which participants organize their activities into "long sequences" [12] of identifiable action types. For example, we have begun to identify sequences in which math problem solving activities are being conducted, as distinct from various other kinds of non-math social interaction. In this way, we are developing a coding scheme that preserves actors' orientations, concerns, relevances and the sequential organization, of the ongoing interaction.This proposed approach to coding makes possible the comparison of different instances of social interaction in ways that preserve the local organization of interaction and exploit that local organization as a source of insight into the ways we come to treat action sequences as sequences of particular sorts.

This paper points the way to achieving an understanding of computer-mediated interaction among multiple participants. As this research has shown, questions concerning the ways that groups are formed and sustained through online interaction can be explored using multiple analytical methodologies as long as care and consideration are given to the differences in the assumptions that inform these methodologies. By using qualitative methods to explain an unexpected analytical result, we have shown how it is possible to interpret the organization of group participation in online interaction.

References

1. Stahl, G. (2002). Rediscovering CSCL. In T. Koschmann, R. Hall & N. Miyake (Eds.), *CSCL 2: Carrying forward the conversation* (pp. 169-181). Hillsdale, NJ: Lawrence Erlbaum Associates. Retrieved from http://www.cis.drexel.edu/faculty/gerry/cscl/papers/ch01.pdf.
2. VMT Project: http://mathforum.org/wiki/VMT/
3. Mason, R. L., Gunst, R. F., Hess, J. L. (2003). *Statistical Design and Analysis of Experiments: With Applications to Engineering and Science, 2nd Edition*. Wiley.
4. Dillenbourg, P., Baker, M., Blaye, A. & O'Malley, C. (1996) The evolution of research on collaborative learning. In E. Spada & P. Reiman (Eds) *Learning in Humans and Machine: Towards an interdisciplinary learning science.* 189-211. Oxford: Elsevier.
5. Avouris, K., Margaritis, F. A Tool to Support Interaction and Collaboration Analysis of Learning Activity, *CSCL 2002*
6. Strijbos, J. W. (2004). *The effect of roles on computer-supported collaborative learning.* Doctoral dissertation, Open University of the Netherlands, The Netherlands.
7. Daradoumis, Th., Martínez, A. and Xhafa, F.: An Integrated Approach for Analysing and Assessing the Performance of Virtual Learning Groups. *CRIWG 2004*: 289-304, Springer
8. Ten Have, P. (1999). *Doing Conversation Analysis: A Practical Guide.* London: Sage Publications.
9. Psathas, G. (1995). *Conversation Analysis: The Study of Talk-in-Interzction.* Thousand Oaks: Sage Publications.
10. Pomerantz, A., Fehr, B. J. (1997). Conversation Analysis: An Approach to the Study of Social Action as Sense Making Practices. In Teun A. Van Dijk (Ed.), *Discourse as Social Interaction, Discourse Studies: A Multidisciplinary Introduction Volume 2,* (pp. 64-91). London: Sage Publications.
11. Heritage, J., Roth A. (1995) Grammar and Institution: Questions and Questioning in the Broadcast News Interview. *Research on Language and Social Interaction*, 28, 1, 1-60.
12. Sacks, H. (1992). *Lectures on Conversation.* Oxford: Blackwell.
13. Livingston, E. (1986). *The ethnomethodological foundations of mathematics.* London, UK: Routledge & Kegan Paul.
14. Kaplan, A. (1964). *The Conduct of Inquiry: Methodology for Behavioral Science.* San Fransisco: Chandler Publishing.