ASSESSMENT INTELLIGENCE IN SMALL GROUP LEARNING

Wanli Xing[1] and Yonghe Wu[2]
[1]*School of Information Science and Learning Technologies, University of Missouri -Columbia, USA*
[2]*School of Open Education and Learning & Shanghai Engineering Research Center of Digital Education Equipment, East China Normal University, Shanghai 20062, China*

ABSTRACT

Assessment of groups in CSCL context is a challenging task fraught with many confounding factors collected and measured. Previous documented studies are by and large summative in nature and some process-oriented methods require time-intensive coding of qualitative data. This study attempts to resolve these problems for teachers to assess groups and give timely feedback. We first operationalize activity theory to holistically frame group work by breaking it down into six dimensions. The captured log data by the collaborative software is mapped with these dimensions and as a result, six measures are generated with a semantic background. Next, we employ a relatively new clustering algorithm – spectral clustering – to categorize groups with similar behaviors, which not only allows us to consider the six indicators simultaneously but also has the capability to deal with large online context. The spectral clustering result is compared with traditional algorithms and demonstrates better assessment accuracy. Furthermore, since the whole process is automated and the group performance indicators are grounded in a meaningful backdrop, it enables teachers to offer concrete and personalized help in a real-time format. Theoretical and practical implications are then discussed.

KEYWORDS

Assessment, CSCL, Clustering, Small Group Learning, Activity Theory

1. INTRODUCTON

Assessment is not only important for evaluating learning outcomes: it may also be a motivating factor for students who have a performance goal orientation (Dennen, 08). While the measurement and assessment of learning is a major responsibility for teachers, it is also a demanding experience for them due to the heavy workload and time-consuming of the assessment activities, especially when learning take place in technology mediated group format (Strijbos, 11; Gress et al., 10). Assessment of group learning in computer-supported environments is more than the measurement of outcomes; the quality of collaborative learning processes is also salient (Strijbos, 11).

However, process-oriented assessment for group learning in the socio-technical context is a difficult problem to resolve in that learning in such an environment takes place through the complex processes and interactions of numerous factors, artifacts, and environments (Barab et al., 01). Because of this complexity, assessment of group remains by-and-large summative in nature (Gress et al., 2010) where the performance of each group was measured by the quality of the solution or product generated. This type of group assessment centers on the intellectual product of the learning process rather than the process itself (Kumar et al., 10), overpassing group dynamics, interaction, and the technology-mediated processes. To address the complex interactions of small groups, some assessment with process-orientation endeavors require integration across multiple coding schemes and sources of data (Hmelo-Silver et al., 11). These assessment efforts often rely heavily on conventional methods such as content analysis, coding of observations, interaction analysis, etc., and are therefore very time-intensive, adding further burdens to the already heavy duties of teachers. Further, most documented studies on group assessment, both summative in nature or towards collaborative process, were conducted after collaboration (Gress et al., 10), and hence are lack of the capability to provide information in a timely format for teachers to act upon.

We describe a methodology that attempts to assess group as the unit in CSCL in order to provide actionable information to teachers in a timely manner. Specifically, this work first explores operationalizing activity theory to holistically frame group activity in a CSCL context by breaking down group work into six dimensions. Then, rather than performing observations or content analysis to produce our measures, electronic trace data generated by the collaborative software are mapped to the six dimensions for measure construction. This step not only lays the groundwork for automating the group assessment, but also contextualizes the data with a semantic grounding. Next, we move beyond identification and analysis of those measures to infer group learning using human judgment but employ a relatively new clustering algorithm to categorize groups with similar behaviors into clusters, which allows us to consider the six indicators simultaneously as well as deal with large online context for offering concrete advice to students. This paper is organized as follows: We first describe the theoretical framework that guides this study. Then the research background is described. Next, methodology, and results are presented. Last, we discuss the implications of our findings for this study as well as pointing out future research directions.

2. THEORETICAL FRAMEWORK

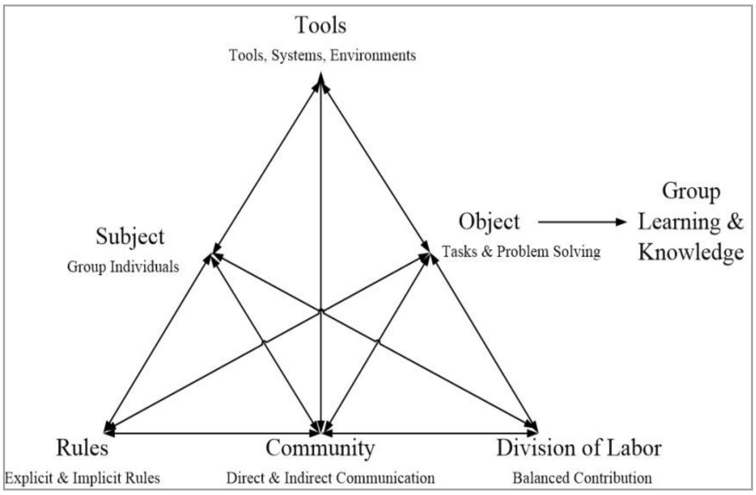

Figure 1. Activity theory in group context

Table 1.Description of activity theory operationalization in CSCL

Measure-metric	Definition
Object	Complete learning tasks collaboratively such as solving a problem or producing an artifact
Subject	Individual student involves and participates in problem solving towards the Object.
Tools	Computers, tools, systems, and environments that mediate the learning and collaboration activity.
Community	Direct and indirect communication enables the group of students to maintain a sense of community and belonging.
Rules	Implicit and explicit rules and guidelines that constrain the activity.
D. of Labor	All group members make concrete and balanced contributions to the overall object.

The Activity System model developed by Engestrom (1987) offers a way to comprehensively frame the collaborative knowledge development process while linking together social behavior and its interdependencies (de Laat, 2006). Engestrom’s (1987) activity model includes six interacting components: subjects, tools, objects, rules, community and division of labor (see Figure. 1). Contradictory tensions between the elements of the Activity System serve to produce the outcome of the activity. The activity of learning (Basharina, 2007) is “the joint activity of a student, physical/symbolic tool(s), and another person(s) performing together as a working social system to achieve some outcome under constraints such as rules.” In our CSCL group assessment context, the outcome and process of this transformation may both be seen as learning and knowledge. It is the sum of the system components and the tensions among them that make up the learning and influence the learning outcomes. Activity Theory helps us to address the complex interactions and see into group performance in a socio-technical CSCL environment (see Table 1).

3. RESEARCH CONTEXT

3.1 Virtual Math Teams

In this study, we operationalize activity theory as a lens for making sense of electronic trace data from a synchronous math tool, Virtual Math Teams with Geogebra (VMTwG) software (Figure. 2). We focused on several modules of a course include teams of three to five practicing teachers who were learning how to implement the curriculum in VMTwG while going through the curriculum themselves. There were 18 groups in total in this course using the VMTwG software. The five modules we analyzed included: “Constructing Dynamic-Geometry Objects”, “Exploring Triangles”, “Creating Construction Tools”, “Constructing Triangles”, and “Inscribing Polygons”. The full curriculum currently includes a total of 21 topics, and is available at the project website (http://vmt.mathforum.org).

Figure. 2 provides us with a visual guide for understanding the cognitive learning discourse in VMT. There are four sections in Figure. 2. Section A reveals the time dimension: the VMT replayer bar. Each action within VMTwG is logged with a timestamp. Section B is the chat window. Here, text is entered in order to chat with others in a group. Future analytics in this project will focus on the automated analysis of the actual text in those windows, in concert with GeoGebra gestures. Sections C and D are related to Geogebra actions. C is the “Take Control” button. Section D is the GeoGebra window itself. Here, students work to create an equilateral triangle within an equilateral triangle using multiple approaches. This is an ordinary part of how VMTwG facilitates interactive problem solving discourse among students. On the other hand, the design of the VMTwG, as a whole, places teams into separate learning spaces for small group interaction. As a result, teachers have difficulty facilitating or directing activity in a timely manner; with 3 or 4 students per group, 18 groups doing an hour of activity translates into 18 hours of replay.

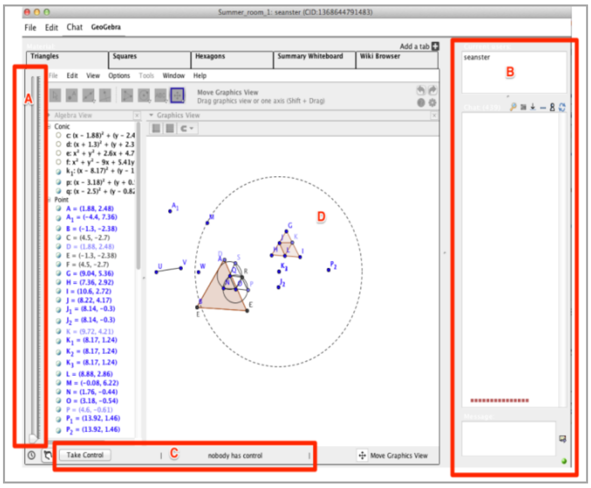

Figure 2.VMTwG of an analytical tool for collaborative math discourse

3.2 Dataset Description

We collected log data in .txt format from five modules in a single course. The data centers on specific event types from the CSCL environment (VMT): Chat, Awareness, Geogebra, System, and WhiteBoard (Wb). The Chat event type logs all the messages that students communicate with each other. Awareness records the actions of erasing the chat messages when the student realizes they are full on the chat bar. Geogebra logs information on how students visually construct a geometry artifact (e.g. add a point, update a segment, etc.). The System event type records information on how the VMT environment is accessed.

4. METHODOLOGY

Based on the trace data captured in the VMT logs, several quantitative measures are developed according to the components of the Activity System as shown in Figure 1. Further details regarding measure construction based on Activity Theory can refer our pervious paper (Withhold for Review).This process of measure construction allows for the development of our automated CSCL assessment technique, while at the same time providing a view of the log data that is intelligible for teachers.

4.1 Measure Construction

4.1.1 Subject

Subject in Activity Theory represents individual students' efforts in solving a problem. As a reflection of individual effort in a group, we consider whether they perform equally in the modules. Therefore, we calculate the frequency of participation under individual tasks and then use the value of standard deviation of these values to represent Subject.

4.1.2 Rules

Rules include implicit and explicit rules. Students have to perform actions that the VMT environment offers. We count the number of distinct functions the group of students used to represent the Rules.

4.1.3 Tools

Under the VMT context, the tools are the System and Wb, where the groups' actions for tool usage are registered. Hence, the Tools dimension for each group is the summation of the actions in these two event types.

4.1.4 Community

All the communications that help maintain the community structure. Community can be presented as the summation value of the frequency of chat messages.

4.1.5 Division of Labor

Division of Labor is a measure of how the balance of the workload is shared among team members. This dimension would have the highest value if all the members in a group shared the workload equally and would have the lowest value if just one of the members took care of the whole problem. Thus, the balance of the work among team members is based on the function group effort (Geogebra dimension) with the perfect division 1/N, where N is the number of students in that Group.

4.1.6 Object

The CSCL activity is to achieve the object of a group of students actively participating in the whole class to solve problems. Object is a function of number of modules involved, event types used, and totally frequency of participation.

In sum, relying on activity theory, we build a quantified model for group performance in CSCL activities specific to the VMT environment: [Subject, Rules, Tools, Community, Division of Labor, Object]. In addition to providing a principled way for measure selection and construction, these theory-grounded measures have a semantic background to facilitate teachers to understand.

4.2 Spectral Clustering

An activity system is characterized by the internal tensions among its components. The tensions are the moving force behind disturbances and innovations and eventually drive the system to change and develop, in this context toward an outcome of group learning. Therefore, it is hard to compute one value as functions of the six dimensions to indicate the learning or performance result of a group of students, especially considering the complexity of the nature of learning. In this exploratory study, we investigated a novel clustering algorithm – spectral clustering – to place groups with similar behavior patterns in the same cluster. Spectral clustering brings into consideration of all the six dimensions in the activity system rather than accounting for only one dimension. Also, it allows teachers to give similar advice to a number of groups at a time, which might be particularly helpful in large-scale online context.

On the other hand, compared with traditional clustering techniques (K-means, EM, etc.), which depend on distances from cluster prototypes usually assuming Gaussian distributions, spectral clustering does not make strong assumptions on the statistics of the clusters but bases on the eigenstructure of an affinity matrix to partition the data objects into disjoint clusters (Von Luxburg, 2007). This feature is important because educational data almost always contains exceptional students and groups (outliers) who can achieve a good performance with little effort or fail without any sensible reason. The cluster statistics does not necessarily follows a certain distribution. Further, empirically, spectral clustering has produced better results than traditional algorithms (Galluccio et al., 2013; Von Luxburg, 2007).

5. RESULTS

5.1 Measures

Table 2.Sample activity theory informed group standardized performance measures

Dimension / Group	Subject	Tools	Community	Rules	Division of Labor	Object
Group 1	0.576	-0.231	-1.066	0.903	0.169	-0.497
Group 2	0.353	-0.376	-0.766	0.721	-0.121	-0.666
Group 3	-0.463	-0.767	0.647	-1.312	0.599	0.724
Group 4	1.550	-0.965	-1.089	1.028	2.091	2.091
Group 5	-1.469	2.247	1.644	-0.550	-0.300	-0.376
…	…	…	…	…	…	…

Group performance is represented as 6 dimension sets in Table 2 (after standardization). There are 18 groups in total. By investigating those numbers alone, the teachers are already able to provide specific actionable advice to a particular group. For example, if the value of a group in the Community dimension (Group 4) is very low, the teacher could suggest the group to communicate more between team members. Similarly, if the group has a low value on Division of Labor (Group 5), then the teacher can tell that the contribution from the group members is not equal. Activity theory equips us with a holistic way to describe groups' participation performance in a CSCL environment rather than via ad-hoc guesswork. These quantified results provide semantic clues that instructors may use to investigate the group work. However, simply examining and comparing those numbers is not a valid or reliable way to assess the group performance. Neither can this method scale up to large online context when there are tens and even hundreds of groups. Therefore, the clustering method from the next section is a good complement to fill these voids.

5.2 Cluster Results

5.2.1 Clustering Performance

Clustering algorithm is performed on the constructed 6-dimension measures informed by Activity Theory. To verify the effectiveness of the clustering result, we used a subjective metric, where the ground truth data is labeled by human evaluators. Specifically, the group performance is judged both by collaborative products and key actions and words used frequency, in which keywords and actions are chosen relying on the authors five years of experience of participation in the VMT development and analysis of small group learning. The list of key actions is: move graphics view, polygon, segment, perpendicular, circle, insert, and triangle, zoom in, erase, join the room. The list of key words and symbols are: right triangle, angle, triangle, equilateral, perpendicular, make, point, control, create, move, drag, scalene, all capital words, ?, !. Then a 1-12 mark is assigned to each student, where 12 is the highest. Then under the circumstance of dataset known classification, a common indicator to assess the clustering results is rand index (RI), a measure that is to test the consistence degree between the clustering results and data external standard classes (Krieger & Green, 1999). The accuracy of rand index is equal to the ratio of the correct matching-pair number to the total of matching-pair number.

$$RI = \frac{\#(the\ correct\ matching\ pair)}{\#(the\ total\ matching\ pair)}.$$

To further validate the proposed spectral clustering method, various cluster number k are chosen k = 2, 3, 4 and also compared the obtained accuracy with the baseline algorithms, k-means and HAC. Accordingly, the ground truth group data are placed into different groups by the concept of percentile in response to different k. For example, if the cluster k is set as 4, then based on the value of the mark labeled, students with 1-3 marks are placed into one group and then 4-6, 7-9 and 10-12 into different groups. Spectral and k-means methods obtained different clusters by directly set the cluster number in the coding process, while HAC got different cluster by choosing different cut standards. Furthermore, with different parameter settings and initializations, 10 different runs are conducted on each algorithm by clustering the constructed 6-diemison measures informed by activity theory. Comparing the generated clusters with the human labeled results, the average values of RI obtained from various algorithms are displayed in Table 3.

To illustrate, the described spectral clustering outperforms the baseline k-means and HAC models in all the conditions of different cluster numbers. It has the best performance when choose cluster k = 2 and reaches 89.4% accuracy.

Table 3. Clustering Performances

Rand Index	Spectral Clustering	k-means	HAC
k = 2	89.4%	83.9%	68.3%
k = 3	85.0%	81.1%	78.3%
k = 4	80.6%	69.4%	56.7%

5.2.2 Cluster Interpretation

In combination with the contextualized measures informed by activity theory, cluster analysis is not only able to facilitate scaling up in online context, but also offer actionable intelligence for the stakeholders. We presented one of the experiments using spectral cluster analysis as an example in Table 4 when we set the cluster number $k=3$. Table 4 shows the standardized means of each cluster over the six dimensions and number of groups in each cluster as well as the range of each dimension. To illustrate the actionable power, Cluster 1 is chosen to demonstrate. When the teacher looks into this table, he or she realize that Cluster 1 has 5 members in it, which tend to have best performances in Tools (0.61), Community (0.46), Rules (0.87) and Object (0.29) dimensions but have lowest ranking in Subject (-0.57) and Division of Labor (-0.45). The low ranking Division of Labor reflects that the members in these groups do not make the same contribution and effort to problem solving. In fact, low value on the Subject aspect is a further proof implying that some of the students in the groups are very active but some are not. The teacher therefore can log into the rooms where those groups are and identify the specific students and encourage balanced contribution to the tasks. Or if in the context of a large number of online groups, the teacher can send one reminder or report to those groups at the same time to remind their problems. While this is from the teacher's perspective, the table is also easily able to present to the students themselves and improve their self-awareness of their performance status.

Table 4.Sample Spectral Cluster k = 3 Results

Cluster Size		Cluster1	Cluster2	Cluster3
		5/18	7/18	6/18
Measure	Range (Standardized)	Cluster Means		
Subject	[-1.47, 2.87]	-0.57	0.45	0.02
Tools	[-1.19, 2.03]	0.61	-0.15	-0.31
Community	[-1.88, 2.09]	0.46	-0.33	-0.33
Rules	[-1.42, 2.25]	0.87	-0.15	-0.75
D. of Labor	[-2.59, 1.10]	-0.45	0.79	-0.35
Object	[-1.67, 2.09]	0.29	0.15	0.08

6. DISCUSSION

The group is the fundamental unit in CSCL. Many studies of group learning use summative assessment methods (e.g. final solution, grade) to measure performance. These assessment approaches usually overlook the collaborative process and the affordances of technology in contribution to group learning (Barab et al., 2001). Qualitative studies e.g. content analysis (Arnold et al., 2009), and conversation analysis (Safin et al., 2010) usually are very time intensive and therefore impractical for teachers to implement. Many quantitative explorations and data mining algorithms usually based on ad-hoc guess work to build their measures and do not address the complex small group dynamics and interaction systematically (Mirriahi et al., 2013). In addition, most of the assessments are conducted after the collaboration and therefore unable to support real-time intervention and reflection. We attempt to solve those problems by designing an automated group assessment methodology that mining the electronic trace data to provide actionable information to teachers in a timely format.

There are multiple perspectives on assessment in the CSCL literature related to group learning e.g. group cognition (Stahl, 2006), learning as participation (Sfard, 1998), and knowledge creation (Lipponen, Hakkarainen & Paavola, 2004). In this paper we zoomed out and explored a different way of developing assessment that is coherent with theory and connected with an advanced algorithm. Activity theory holistically frames group participation in our framework, which addresses situatedness, contextuality, and social mediation. From a practical perspective, there is an overreliance on text-based measures to assess learning in CSCL (Gress et al., 2010). Coding of the discussions and content are usually quite time-consuming. It is extremely difficult for a real world teacher to code and provide timely feedback to students. In addition, these coding methods and frameworks are not always shared (Gress et al., 2010) which leads to difficulty in maintaining consistency among different evaluators. The proposed method, from measure construction to clustering is totally automated. Therefore, it could significantly reduce teachers' assessment burden and has the affordance for teachers to provide timely feedback. The automated method can also increase the consistency of evaluations and improve the reliability of the results.

7. CONCLUSION

This study presents an activity theory-grounded spectral clustering model for assessing groups in a technology-mediated environment. We first operationalized activity theory to holistically quantify the group participation in VMT environment and generated 6 measures. Then, the spectral clustering algorithm is coded and the assessment accuracy is compared with traditional techniques, which the proposed spectral clustering achieved the best accuracy. Since the whole process is automated and the generated group performance indicators are with a meaningful semantic background, it allows teachers to provide concrete and personalized feedback in a real-time manner. A combination of qualitative and quantitative assessment measures may offer the best way to assess learning and performance in a CSCL environment. However, qualitative assessment is often time consuming. Future work can investigate the technique of natural language processing of the chat log data and incorporate the results into the activity theory measure construction system in order to further inform group assessment in CSCL.

REFERENCES

Arnold, N., Ducate, L., Lomicka, L., & Lord, G. (2009). Assessing online collaboration among language teachers: A cross-institutional case study.Journal of Interactive Online Learning, 8(2), 121-139.

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

Dennen, V. P. (2008). Looking for evidence of learning: Assessment and analysis methods for online discourse. Computers in Human Behavior, 24(2), 205-219.

Engeström, Y. (1987). Learning by expanding: An activity-theoretical approach to developmental research.

Galluccio, L., Michel, O., Comon, P., Kliger, M., & Hero, A. O. (2013). Clustering with a new distance measure based on a dual-rooted tree.Information Sciences, 251, 96-113.

Goldstone, R. L., & Wilensky, U. (2008). Promoting transfer by grounding complex systems principles. The Journal of the Learning Sciences, 17(4), 465-516.

Gress, C. L., Fior, M., Hadwin, A. F., & Winne, P. H. (2010). Measurement and assessment in computer-supported collaborative learning. Computers in Human Behavior, 26(5), 806-814.

Halverson, C. A. (2002). Activity theory and distributed cognition: Or what does CSCW need to DO with theories?. Computer Supported Cooperative Work (CSCW), 11(1-2), 243-267.

Hmelo-Silver, C. E., Jordan, R., Liu, L., & Chernobilsky, E. (2011). Representational tools for understanding complex computer-supported collaborative learning environments. In Analyzing Interactions in CSCL (pp. 83-106). Springer US.

Krieger, A. M., & Green, P. E. (1999). A generalized Rand-index method for consensus clustering of separate partitions of the same data base. Journal of classification, 16(1), 63-89.

Kumar, V. S., Gress, C. L., Hadwin, A. F., & Winne, P. H. (2010). Assessing process in CSCL: An ontological approach. Computers in Human Behavior, 26(5), 825-834.

Leont'ev, A. N. (1974). The problem of activity in psychology. Journal of Russian and East European Psychology, 13(2), 4-33.

Lipponen, L., Hakkarainen, K. & Paavola, S. (2004). "Practices and Orientations of CSCL," *What We Know about CSCL: And Implementing It in Higher Education* (pp. 31-50), Kluwer Academic/Springer, Boston, US.

M.F. de Laat (2006). "Networked Learning", *PhD dissertation*, Utrecht Univ., the Netherlands.

Mirriahi, N., & Dawson, S. (2013, April). The pairing of lecture recording data with assessment scores: a method of discovering pedagogical impact. In*Proceedings of the Third International Conference on Learning Analytics and Knowledge* (pp. 180-184). ACM.

Safin, S., Verschuere, A., Burkhardt, J. M., Détienne, F., & Hébert, A. M. (2010). Quality of collaboration in a distant collaborative architectural educational setting. *International Reports on Socio-Informatics*, *7*(1).

Sfard, A. (1998). On two metaphors for learning and the dangers of choosing just one. *Educational researcher*, *27*(2), 4-13.

Strijbos, J. W. (2011). Assessment of (computer-supported) collaborative learning. *Learning Technologies, IEEE Transactions on*, *4*(1), 59-73.

Von Luxburg, U. (2007). A tutorial on spectral clustering. *Statistics and computing*, *17*(4), 395-416.