Xing, W., Wadholm, R., Petakovic, E., & Goggins, S. (2014). Group Learning Assessment: Developing a Theory-Informed Analytics. *Educational Technology & Society*, *xx* (x), xx–xx.

Group Learning Assessment: Developing a Theory-Informed Analytics

Wanli Xing, Robert Wadholm, Eva Petakovic and Sean Goggins*
University of Missouri, Columbia, MO, 65203, USA // wxdg5@mail.missouri.edu // rrwb84@mail.missouri.edu// ehpppf@mail.missouri.edu // GogginsS@missouri.edu

*Corresponding author

ABSTRACT

Assessment in Computer Supported Collaborative Learning (CSCL) is an implicit issue, and most assessments are summative in nature. Process-oriented methods of assessment can vary significantly in their indicators and typically only partially address the complexity of group learning. Moreover, the majority of these assessment methods require time-intensive coding of qualitative data. Our study explores the operationalization of activity theory to frame group activity in a CSCL context by breaking group work into six dimensions. We map log data generated by a collaborative software, Virtual Math Teams with Geogebra, with these dimensions and construct six measures to lay the groundwork for automating CSCL assessment. Next, we move beyond identification and analysis of those measures to infer group learning using human judgment and employ a clustering algorithm to categorize groups with similar performance, allowing us to consider the six indicators simultaneously and to step further toward assessment automation. Last, in terms of the complexity of group learning in a socio-technical context, we discuss a web-based tool that not only shows group-level assessment but also integrates with our previous work on individual assessment, thus providing teachers with a diverse of group learning.

Keywords

Assessment, CSCL, Activity theory, Cluster analysis, Automation

Introduction

A growing body of research has coalesced around the idea that groups, not individuals, are the principle engines of learning (Strijbos, 2011; Stahl, 2006). According to Vygotsky (1978), high-level cognitive functions appear first as interpsychological processes, or group learning, and only later as intrapsychological learning, which results from the internalization of social participation. Learning in groups is a matter of participation in social processes and interactions with artifacts (Lipponen, Rahikainen, Lallimo & Hakkarainen, 2003). According to Stahl (2006), group cognition presupposes three levels of learning: individuals, small groups, and communities. Different analyses and assessment methods may occur at each level, and these levels influence each other and are best viewed as an integrated, complex whole: understanding a group in its context becomes a key to measuring learning individually and at the community level.

However, assessment of in-group learning remains by-and-large summative in nature (Gress, Fior, Hadwin & Winne, 2010): the qualities of group outcomes or collaboration products are considered the key criterion for assessing collaborative learning. While "real-time," "process-orientation" and "socio-technical context" are key characteristics of the group concept in collaborative learning (Sfard, 1998; Reimann, 2009), summative evaluations are usually administered after the collaboration, which fundamentally undermines the theoretical constructs behind group learning in CSCL. Assessments conducted during collaboration have various constructs and interests and usually address part of group dynamics and collaboration for learning. On the other hand, a majority of assessment methods during collaboration have consisted of observations and content and interaction analyses (Strijbos, 2011; Gress et al, 2010). Coding these kinds of data is usually time-intensive (Daradoumis et al, 2006). It is difficult, if not impossible, to implement this kind of group assessment on a large scale.

Our study explores operationalizing activity theory to frame group activity in a CSCL context by breaking down group work into six dimensions. Then, rather than performing observations or content analysis to generate our measures, we construct measures based on electronic trace data generated by collaborative software. Next, we move beyond identification and analysis of those measures to infer group learning using human judgment, but employ a clustering algorithm to categorize groups with similar participation performance and further automate assessment. Last, in terms of the complexity of collaborative learning in a socio-technical context, we discuss how a web-based tool that is in development integrates with our previous work on individual assessment, providing teachers with a holistic and multi-perspective view of group learning.

Literature Review

The activity of assessment can strongly influence learning (Frederiksen, 1984). Research on collaborative learning assessment has largely followed two lines (Gress et al, 2010): assessing the individual and assessing the group. Although studies diverge along these two lines, Sfard (1998) proposes a theoretical unification of collaborative learning and assessment, arguing that to focus only on individuals or groups may lead to distortions and undesirable practices. Xing, Wadholm & Goggins (2014a) conducted an extensive review of assessment methods and analysis at the individual level. The authors constructed an activity theoretic, quantified model consisting of six measures to holistically gauge individual participation in group learning activities based on electronic trace data. Teachers can then use these constructed measures to identify and provide feedback regarding the shortcomings of students.

Assessment of groups in collaborative learning has been dominated by “after collaboration” measurement (Gress et al., 2010). After reviewing 186 articles, Gress et al. state that over half of studies measuring collaborative learning were conducted after collaboration. Typically, groups in CSCL are evaluated through the assessment of final products. Kapur and Kinzer (2009) looked at productive failure in CSCL by comparing groups solving ill- and well-structured problems. The performance of each group was measured by the quality of the solution produced, using a rubric for assessment.

Assessments performed during collaboration have a variety of foci. Gress et al. (2010) found 71 instances of metrics gauging group collaboration in four categories derived from prior literature: first, the investigation of successful collaboration. For instance, Safin et al. (2010) video-recorded the collaborative activities of two teams in the domain of architectural design, assessing for successful collaboration based on their own protocol developed from seven dimensions including fluidity of collaboration, task and time management, and individual task orientation. Arnold, Ducate, Lomicka and Lord (2009) studied foreign language graduate student collaboration in a wiki context. Through content analysis of the dialogue with reference to Curtis and Lawsons’ Framework, they constructed five measures to determine successful student collaboration.

The second construct focuses on group performance using GeoGebra. This builds on the work of Fessakis et al. (2013), who developed a graphical interaction analysis tool to investigate activity levels within and among groups in a blog environment by counting connections with other members and number of posts. Their results show a significant impact of the graph’s presence on group interaction. The interaction graph generation process used for assessment was semi-automated in this research. Our third research construct focuses on social interactions and communication within collaborative settings. We build on Calvani et al. (2010), who proposed eleven indicators to measure the effect of interaction. Their group indicators relied on content analysis of group dialogue and on message counts. Similarly, De Laat et al. (2006) used social network theory to study patterns of communication in a team.

The last area of focus of group assessment is knowledge construction and skill acquisition. Kumar et al. (2010) built a tool to inspect group knowledge construction and cognition using an ontological approach. They tried to consider the full range of collaborative activities that relate to group learning. The central question is “How does one go about designing a system that affords multiple collaborative opportunities while capturing all learner activities and the full context of their learning environments (e.g., open documents, notes being made, searches, conversations, and sharing) all on a standardized underlying metric…?” They built their own ontologies to assess those aspects based on electronic trace data. Milrad (2002) employed construction kits, modeling tools, and system dynamics simulations to inquire into collaborative discovery learning. All collaborative activities were recorded and transcribed to assess students’ skills and knowledge acquisition.

In summary, group assessment in CSCL mainly takes place after collaboration and focuses on the collaborative product rather than the process. Due to the diversity of process-oriented measures for group assessment in the CSCL community, research explorations vary widely with different interests and constructs covering artifacts, context, interactions, and knowledge development. However, few studies are able to evaluate groups in CSCL holistically. In addition, assessments with process-related metrics rely too heavily on conventional methods such as content analysis, coding of observations and video recordings. These measures provide invaluable information, but since this analysis is very time-intensive and is conducted post-event, information obtained does not offer opportunities for real-time feedback to enhance evaluation, reflection, awareness and adaptation (Kumar et al. 2010). Though there are different perspectives on group learning, the most common ones share a focus on participation as a condition for learning. Our method approaches group assessment from an activity theory perspective and centers on participation.

Theoretical Framework

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). To illustrate, an activity system provides three characteristics for analyzing learning in group work (De Laat, 2006): activity theory focuses on contextuality and is oriented towards systematically comprehending group dynamics, objects, mediating artifacts and social organization; it also relies on a dialogical theory of knowledge and thinking (language and communication) with a concentration on human cognition; further it is a developmental theory aiming to explain changes in human practices over time.

Engestrom’s (1987) activity model includes six interacting components: *Subjects*, *Rules*, *Tools*, *Community*, *Division of Labor*, and *Object* (see Fig. 1). The activity of learning 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 process and outcome 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 learning and influence learning outcomes.

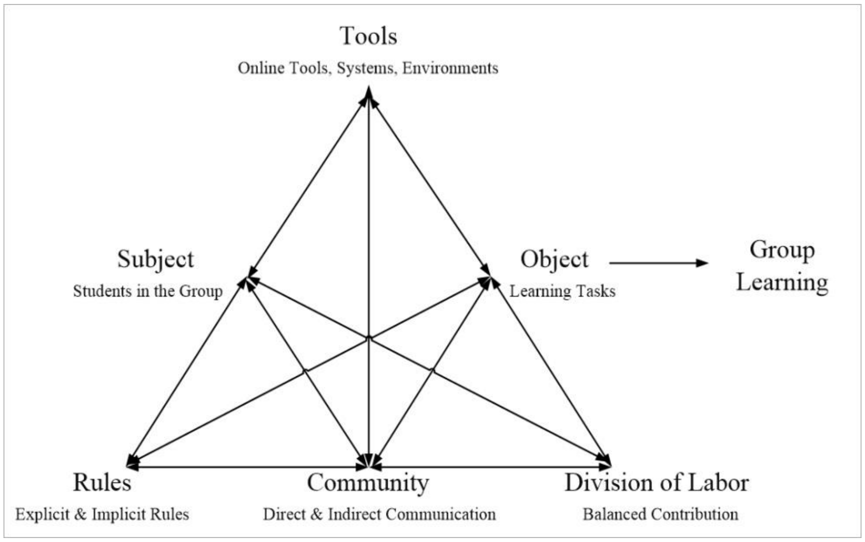


Figure 1. Activity theory for a group

Current assessment practices of group performance often address only part of the activity of the learning system, focusing on process, context, cognition, artifacts or a mix of two or three such components. Activity theory allows us to address the complex interactions and envision group performance in a socio-technical CSCL environment (see Table 1).

Table 1. Description of activity theory operationalization

Measure-Metric	Definition
Subject	Individual student effort that contributes to problem solving.
Rules	Implicit and explicit rules and guidelines constraining the activity. For example, teachers can set specific rules for a learning task (explicit), and students can use only the functions supported by the available tools and are bounded by socially established (implicit) norms.
Tools	Computers, online tools, systems, and environments that mediate the learning and collaboration activities.
Community	Direct and indirect communication enabling the group of students to maintain a sense of community and belonging
Division of Labor	Overall coordination among group members.
Object	Jointly completed learning tasks such as the solution of a problem or production of an artifact (e.g. essays).

Research Context

VMT

In this study, we operationalize activity theory in order to make sense of electronic trace data to assess groups using a synchronous math discourse tool. We focus on several modules of a 2013 course designed for Virtual Math Teams with Geogebra (VMTwG) software (Figure 2).

The five modules analyzed include teams with 3-5 members who were going through the curriculum using VMTwG. Fourteen groups were chosen for this study. The 5 modules that were analyzed included: “Construcing 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 a guide to understanding cognitive learning discourse in VMT. Section A of Figure 2, the VMT replayer bar, represents the time dimension. Each action within VMTwG is logged with a timestamp. Section B is the chat window, where text is entered into chat. Future analytics in this project will focus on the analysis of the text in chat windows, in concert with GeoGebra gestures. Sections C and D are related to Geogebra actions. C is the “Take Control” button, which gives an individual user control of the tools. Section D is the GeoGebra window. Here, students use multiple approaches to inscribe an equilateral triangle within another equilateral triangle.

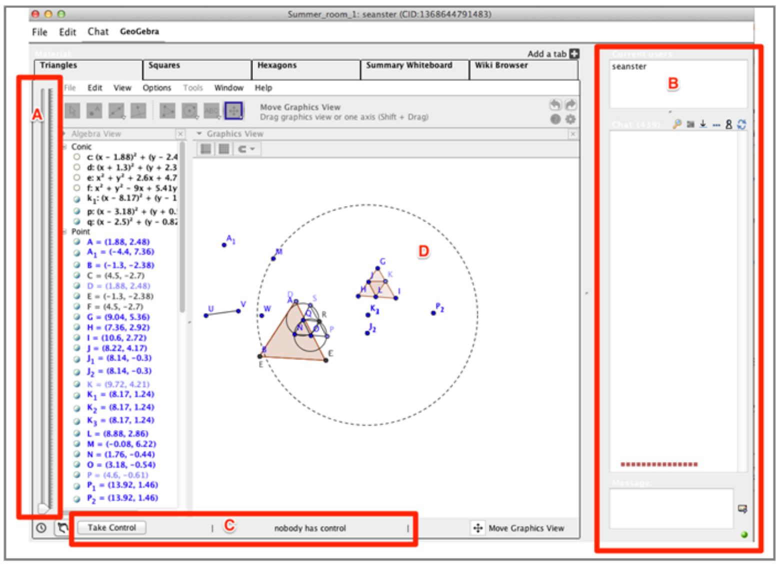


Figure 2. VMTwG, an analytical tool for collaborative math discourse

Dataset Description

Community	Subject	Topic	Room	Source	Target	Time	Finish Time	Event Type	Event
Spring 2013	Dynamic Geometry	Topic 03	Holland_Group_4	madison_	emma_r	0:00:48	26:52.8	chat	2013-03-08 13:26:52.799 - madison_m -> what do we do now im confused what tab are we in??
Spring 2013	Dynamic Geometry	Topic 03	Holland_Group_4	emma_r	madison_	0:0:-44	26:08.6	Geogebra:I	2013-03-08 13:26:08.643 - emma_r -> added line:Line i
Spring 2013	Dynamic Geometry	Topic 03	Holland_Group_4	emma_r	emma_r	0:00:00	26:13.0	awareness	2013-03-08 13:26:12.987 - emma_r -> [fully erased the chat message]
Spring 2013	Dynamic Geometry	Topic 03	Holland_Group_4	emma_r	emma_r	0:00:17	26:31.6	chat	2013-03-08 13:26:31.614 - emma_r -> i just made the line segment between my new points
Spring 2013	Dynamic Geometry	Topic 03	Holland_Group_4	amina_p	emma_r	0:00:00	26:40.7	chat	2013-03-08 13:26:40.697 - amina_p -> k
Spring 2013	Dynamic Geometry	Topic 03	Holland_Group_4	emma_r	amina_p	0:00:23	27:04.0	Geogebra:I	2013-03-08 13:27:03.997 - emma_r -> added line:Line j
Spring 2013	Dynamic Geometry	Topic 03	Holland_Group_4	emma_r	emma_r	0:00:03	27:15.0	chat	2013-03-08 13:27:15.045 - emma_r -> bisector
Spring 2013	Dynamic Geometry	Topic 03	Holland_Group_4	emma_r	emma_r	0:00:21	28:03.7	chat	2013-03-08 13:28:03.744 - emma_r -> no i cant construct the other line segments. maybe you guys can
Spring 2013	Dynamic Geometry	Topic 03	Holland_Group_4	emma_r	emma_r	0:00:00	28:04.1	Geogebra:I	2013-03-08 13:28:04.139 - emma_r -> tool changed to Move Graphics View
Spring 2013	Dynamic Geometry	Topic 03	Holland_Group_4	emma_r	emma_r	0:00:06	28:21.2	chat	2013-03-08 13:28:21.155 - emma_r -> who wants control now?
Spring 2013	Dynamic Geometry	Topic 03	Holland_Group_4	madison_	emma_r	0:00:12	28:33.9	system	2013-03-08 13:28:33.946 - madison_m -> Now viewing tab TEAM 4 TAB
Spring 2013	Dynamic Geometry	Topic 03	Holland_Group_4	madison_	madison_	0:00:03	28:37.3	system	2013-03-08 13:28:37.313 - madison_m -> Now viewing tab Bisector
Spring 2013	Dynamic Geometry	Topic 03	Holland_Group_4	madison_	madison_	0:00:19	29:01.9	chat	2013-03-08 13:29:01.877 - madison_m -> ive added a tab its called....TEAM 4 TAB
Spring 2013	Dynamic Geometry	Topic 03	Holland_Group_4	emma_r	madison_	0:00:08	28:52.3	chat	2013-03-08 13:28:52.349 - emma_r -> why did you add a new tab madison>
Spring 2013	Dynamic Geometry	Topic 03	Holland_Group_4	emma_r	emma_r	0:00:01	28:54.8	chat	2013-03-08 13:28:54.817 - emma_r -> ?
Spring 2013	Dynamic Geometry	Topic 03	Holland_Group_4	emma_r	emma_r	0:00:00	28:57.3	chat	2013-03-08 13:28:57.335 - emma_r -> ?
Spring 2013	Dynamic Geometry	Topic 03	Holland_Group_4	emma_r	emma_r	0:00:02	28:59.6	system	2013-03-08 13:28:59.602 - emma_r -> Now viewing tab TEAM 4 TAB
Spring 2013	Dynamic Geometry	Topic 03	Holland_Group_4	emma_r	emma_r	0:00:16	29:16.2	wb	2013-03-08 13:29:16.190 - emma_r -> emma_r created a scribble
Spring 2013	Dynamic Geometry	Topic 03	Holland_Group_4	emma_r	emma_r	0:00:00	29:16.6	system	2013-03-08 13:29:16.629 - emma_r -> Now viewing tab Bisector
Spring 2013	Dynamic Geometry	Topic 03	Holland_Group_4	emma_r	emma_r	0:00:02	29:18.7	Geogebra:I	2013-03-08 13:29:18.654 - emma_r -> tool changed to Move
Spring 2013	Dynamic Geometry	Topic 03	Holland_Group_4	madison_	emma_r	0:00:34	30:04.0	chat	2013-03-08 13:30:03.979 - madison_m -> i feel abandond whats wrong guys?????? ;(
Spring 2013	Dynamic Geometry	Topic 03	Holland_Group_4	amina_p	madison_	0:00:22	29:55.0	chat	2013-03-08 13:29:55.023 - amina_p -> are finish with the instructions emma
Spring 2013	Dynamic Geometry	Topic 03	Holland_Group_4	emma_r	amina_p	0:00:00	29:48.5	chat	2013-03-08 13:29:48.485 - emma_r -> i get it
Spring 2013	Dynamic Geometry	Topic 03	Holland_Group_4	emma_r	emma_r	0:00:01	29:52.6	chat	2013-03-08 13:29:52.573 - emma_r -> geeeezzz
Spring 2013	Dynamic Geometry	Topic 03	Holland_Group_4	emma_r	emma_r	0:00:15	30:08.1	Geogebra:I	2013-03-08 13:30:08.068 - emma_r -> tool changed to Intersect Two Objects

Figure 3. Sample logs from VMT

All log data for this study was collected in .txt format, which centers on specific event types: VMT Awareness, Geogebra, System, Chat, and WhiteBoard (Wb). The Chat event type logs all the messages in the group. Awareness records the action of erasing chat messages. Geogebra logs information on how students virtually construct a geometry artifact (adding a point, updating a segment). The system logs all user actions, and automates the construction of action adjacencies to reveal the order each participant works in. Figure 3 shows a sample of original log data.

Methodology

Data Categorization

Our analysis examines each interaction as related to four dimensions for a single student in each group: *Individual, Group, Action Types, and Module*. The *Individual* category sums all the personal actions in every event type for each individual member of the group (frequency in each event type) in the situation where the student is both the creator (source) and the receiver (target) of the action according to Figure 3. Similarly, the *Group* category sums all actions taking place between a student and their team members (where the student is the creator and the team member is the receiver). The *Action Types* dimension is a set of values containing the kinds of actions over that particular event type. For example, if a student never erased a message in the Awareness event over all the modules, then the action types for Awareness is {0}; for a Wb event, if a student takes actions such as creating a textbox and copying an object, but never uses other actions such as moving objects, resizing, etc., then the action types in Wb are {create a textbox, copy an object}; for the System event, a student's action constraints might be represented as {Join a room, View a Tab}. The awareness dimension is noted according to curricular module: {Module 1, Module 3, Module 4}.

Data Formation

Log data is then processed to center on the group as a unit. *m* notates the group, $m=1,2,\cdots,M$. In this context, M is 14 indicating that there are 14 groups in total. N represents the number of students in that particular Group *m*, $n=1,2,\cdots,N$. N may be different for different groups. a

represents a single student. T denotes the event type, T_{ij} represents the value of measurement aspects j in event type i, where $i \in [1,5]$, $j \in [1,4]$, $i \in Z^+$, $j \in Z^+$ .To be more specific, there are five event types(i.e., Awareness, Geogebra, System, Chat, and Wb) and four measurement aspects (i.e., *Individual, Group, Action Types, Module*). *G* denotes the 6 dimensions in activity theory for each group. G_{1m} represents the m^{th} group's performance in the dimension of *Subject*, G_{2m} notates the m^{th} group's performance in the dimension of *Rules*, and so on. For the purpose of this study, all data will focus on the group unit of analysis. In combination with the *Data Categorization* section, the data set for a single group in this course is a four-fold structure shown in Figure 4:

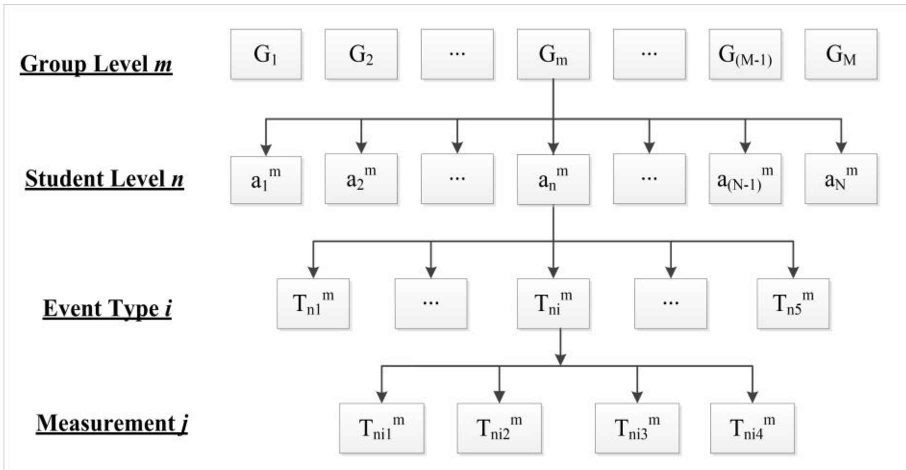


Figure 4. Four-fold data structre

Measure Construction

Subject

Subject in activity theory represents the individual student's efforts in solving a problem. When mapped to our log data, individual endeavor is a student's actions over the five event types across all modules where he or she is the starter and receiver. Therefore, for a single student in a group, individual effort can be quantified as:

$$a_{n1}^{m} = \sum_{i=1}^{5} T_{ni1}^{m} \ .$$

Because the *Subject* dimension in a group is still noted as the individual aspect of endeavor, subjective efforts can first be considered as the accumulation of individual effort from the group members.

$$A_1^m = \frac{1}{N}\sum_{n=1}^{N} a_{n1}^{m} = \frac{1}{N}\sum_{n=1}^{N}\sum_{i=1}^{5} T_{ni1}^{m} \ .$$

The average for each group is used here because groups may vary in number of students. Otherwise, groups with more students would have a much higher chance to outperform groups with fewer students.

As a reflection of individual effort in a group, we need to consider whether they perform equally in the modules. Therefore, we use the standard deviation of the group members to show equal contributions:

$$SD_1^m = \sqrt{\frac{1}{N}\sum_{n=1}^{N}(a_{n1}^m - \frac{A_1^m}{N})^2} .$$

Therefore, *Subject* can be denoted as:

$$G_{m1} = \frac{A_1^m}{SD_1^m} .$$

Rules

Rules may be implicit or explicit. Considering that in this course the students conducted themselves through the curriculum, explicit rules have no obvious effect on the group. Under the socio-technological construct, the rules are the implicit rules that constrain students' actions, denoted as T_{ni3}. Students can only perform actions that the VMT environment offers.

T_{ni3} is a set that contains action types in a particular event type. To consider the group performance in Rules aspect, we process group performance over this dimension in each event type separately. First, we put the student members' *Action Types* over one event type in a unified set. Then for a particular kind of action, for example, "View a Tab" under System, if the students in a group all perform the action, we noted as the number $K_{i3}=N$. Following this logic, if *N-1* of *N* students used this "View a Tab" action, it is denoted as $K_{i3}=N-1$. Therefore, for each rule in each event type, the range $K_{i3} \in [0, N]$.

On the other hand, each event type has a different number of rules. To illustrate, Awareness may only contain one rule (Erase the Message) while System has three (Join the Room, Leave the Room and View a Tab). Therefore, combined rule sets over a particular event type may result in a different number. We use $L = max(|T_{ni3}|)$ and $n = 1,2,\cdots,N$ to represent all possible rules contained in a particular event type. Therefore, for any event type T, the number of students using one particular rule can be presented as: $\{K_{i3}^1, K_{i3}^2, ..., K_{i3}^L\}$

A simple sum of those numbers could reflect group performance over the rule aspect to some extent. However, it is difficult to differentiate performance among groups because the scale of this summation is too small. For another reason, intuitively, a group of four or five all performing one kind of action requires more effort than a group of three. Thus, we also need to expand the difference among groups with different numbers of group members. We use e as the base to achieve this goal. Therefore, *Rules* is presented as:

$$G_{m2} = \frac{1}{N}\sum_{i=1}^{5}(e^{K_{i3}^1} + e^{K_{i3}^2} + \cdots + e^{K_{i3}^L}) .$$

Tools

The *Tools* dimension in activity theory focuses on the process in which tools facilitate group knowledge development. In the VMT context, the tools are System and Wb, where the groups' actions for tool usage are registered, denoted as T_{ni1} and T_{ni2}. Hence, for Group m, the activity theory *Tools* dimension G_{m3} can be presented as:

$$G_{m3} = \frac{1}{N}\sum_{n=1}^{N}\sum_{i=4}^{5}(T_{ni1}^{m} + T_{ni2}^{m}) .$$

Community

Community can be presented as the sum of chat messages and awareness. In the VMT context, students in the same group use chat to directly communicate with each other, and use the awareness function to erase the chat messages which can be categorized as an indirect contribution to the community, denoted as T_{ni1} and T_{ni2}. Then, for group m, the activity theory *Community* dimension G_{4m} can be denoted as:

$$G_{m4} = \frac{1}{N}\sum_{n=1}^{N}\sum_{i=1}^{2}(T_{ni1}^{m} + T_{ni2}^{m}) .$$

Division of Labor

Division of Labor is a measure of how workload is shared among team members. Labor consists of contributions made by the student to the collaborative learning modules. Although chat messages may also contribute to the development of geometry objects, all concrete contributions to geometry object construction originate from the Geogebra dimension to build the actual geometry objects. Therefore, labor of each student can be denoted as:

$$a_{n5}^{m} = \sum_{i=1}^{5}(T_{in1}^{m} + T_{in2}^{m}) .$$

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 only one member did the work. Thus, the balance of the work among team members is based on the standard deviation of the group effort with the perfect division $\frac{1}{N}$. However, the smaller the standard deviation value is, the more balanced the effort is among team members. This trend is contrary to the other 5 dimensions. Therefore, we use a minus value to change the trend. In sum, the *Division of Labor* can be represented as:

$$G_{m5} = 1 - \sqrt{\frac{1}{N}\left[\left(\frac{a_{15}^{m}}{\sum_{n=1}^{N} a_{n5}^{m}} - \frac{1}{N}\right)^2 + \left(\frac{a_{25}^{m}}{\sum_{n=1}^{N} a_{n5}^{m}} - \frac{1}{N}\right)^2 + \cdots + \left(\frac{a_{N5}^{m}}{\sum_{n=1}^{N} a_{n5}^{m}} - \frac{1}{N}\right)^2\right]} .$$

Object

The CSCL activity is to achieve the object of a group of students actively participating in the whole class to solve problems. Hence, the first factor to consider is the number of modules in which the group participates. In order to quantify whether the group is active in those learning modules, we incorporate the total frequency of participation and the number of event types. By doing this, we avoid inflated ratings for groups who participate in all modules but make few actions or contributions.

Similar to the Rules function, T_{ni4} is a set containing modules in which students participate in that particular event type. Involved in this module is learning object construction. Therefore, instead of a union set ($\cup$) of all the modules students participate in within a particular event type, we implement a harsher standard intersection ($\cap$) to highlight its significance. Therefore, for each event type, unless all members in that group are involved in that single module, that module is not counted as a valid number to compute group performance over the object dimension. Then the sum of modules in which all students are involved in over five event types is counted as one aspect of Object.

On the other hand, the number of modules the group is involved in is the key for the *Object* dimension; the other two are secondary factors. While the frequency for participation is in the 1000s, the number of modules is in the scale of 10s, so we use the log function to dampen the effect of the frequency measure. Although the event types are in the same scale of modules, we still want to dial down for overall event types that the group used. Thus, we use a fraction to lower the effect of event type on object measurement. As a result, each group on *Object* aspect can be denoted as:

$$G_{\mathrm{m6}} = [\frac{1}{N}\sum_{n=1}^{N} log \sum_{i=1}^{5}(T_{ni1}^{m} + T_{ni2}^{m})]*[\frac{1}{N}\sum_{\mathrm{i}=1}^{5}\left|T_{1i4} \cap T_{2i4} \cap \cdots T_{ni4} \cap \cdots \cap T_{Ni4}\right|]*[\frac{1}{N}\sum_{n=1}^{N}(\left|T_{ni}\right| / 5)]$$

In sum, based on activity theory, we have built a quantified model for group performance in CSCL activities specific to the VMT environment as shown in Figure 5: [*Subject, Rules, Tools, Community, Division of Labor, Object*].

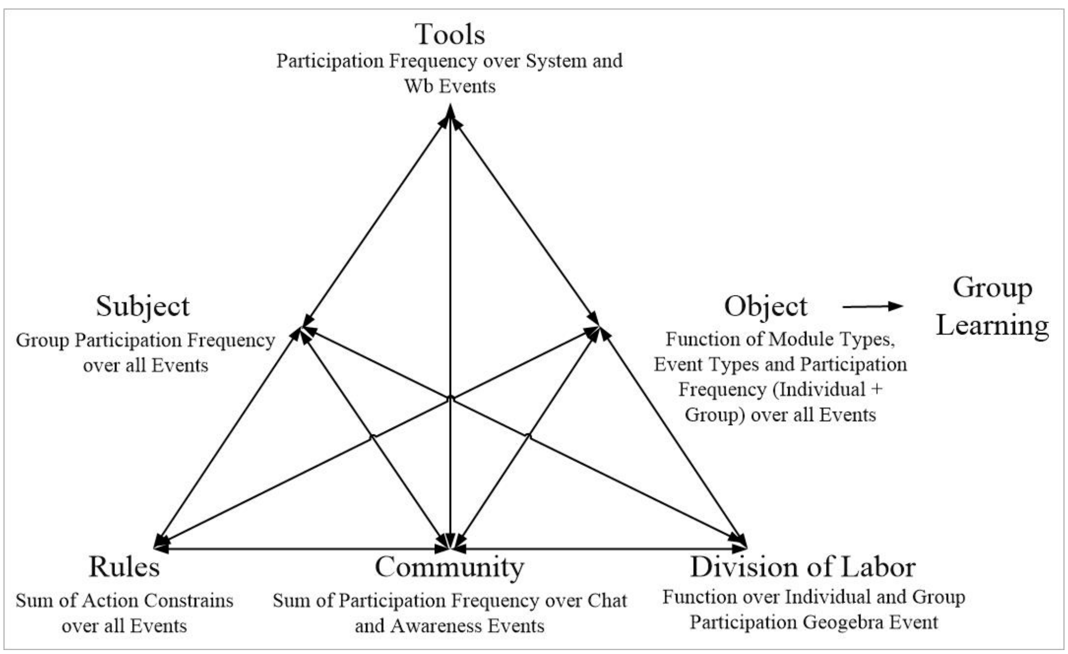


Figure 5. Quantified activity theory model for a group

Group Assessment of Participation in VMT

Rational

An activity system is characterized by the internal tensions among its components that eventually drive the system to change and develop, in this context, toward an outcome of group learning. Therefore, it is difficult 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 group learning in a technology-mediated environment. In this exploratory study, we use K-means cluster analysis to chunk groups with similar learning behaviors. A K-means cluster brings into consideration each of the six dimensions in the activity system rather than accounting for only one dimension. Another advantage of using cluster analysis is to enable the teacher to understand the overall collaborative activity performance of the whole class.

Procedure

Cluster analysis, which addresses the problem of data segmentation, belongs to unsupervised learning methods, as there is no knowledge of "preferred" clusters (Guo & Zhang, 2014). It is a set of techniques used to classify a data set into groups that are relatively homogeneous within themselves and heterogeneous between each other on the basis of a defined set of variables (Xing, Guo, Fitzgerald, & Xu, 2014). A significant step in clustering is to define the system scale and select the proper cluster elements, which we have defined earlier. By considering the measurements defined in the subsection of measure construction, it is possible to group students into different categories with a mathematical method. Therefore, the state definition used for this study is a vector of measurements for all the groups. The data samples are therefore multidimensional because the vectors of measures for each group are considered simultaneously (six dimensions developed from activity theory). Hence, the system states in our study are defined as follows, assuming there are 6 representative measurements in the datasets and *N groups*. For instance, the constructed measure can be recorded for N individuals (groups). Then the data A, a $K(6)\times K(N)$ matrix, will have the format as follows:

$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1m} & \cdots & x_{15} & x_{16} \\ x_{21} & x_{22} & \cdots & x_{2m} & \cdots & x_{25} & x_{26} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nm} & \cdots & x_{n5} & x_{n6} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{(N-1)1} & x_{(N-1)2} & \cdots & x_{(N-1)m} & \cdots & x_{(N-1)5} & x_{(N-1)6} \\ x_{N1} & x_{N2} & \cdots & x_{Nm} & \cdots & x_{N5} & x_{N6} \end{bmatrix}$$

To manage differences in scale among variables within each group, the cluster elements should be properly normalized. This process, which uses Eq. (1), is performed before the cluster analysis to reduce dimensionality of the original data.

$$x'_{nm} = \frac{x_{nm} - \bar{x}_m}{s_m}, \quad (n=1,2,\cdots,N; \quad m=1,2,\cdots,M)$$

In cluster analysis, cluster elements are grouped according to their similarities, or more specifically, the distances between them (Xing, Guo, Lowrance & Kochtanek, 2014). Smaller distances between elements indicate greater similarity and higher likelihood of belonging to the same cluster. For our study, squared Euclidean distance, as shown in Eq. (2), is implemented to calculate distance between clusters.

$$D_{ij}^2 = \sum_{m=1}^{M} (x_{im} - x_{jm})^2, \quad (i=1,2,\cdots,N; \quad m=1,2,\cdots,M)$$

Where, D^2_{pq} is the squared Euclidean distance between measurements p and q ; x_{im} is the m^{th} element in individual *i*; and x_{jm} is the m^{th} element in individual *j*.

Results

Activity Theory based Measurements

Group performance is represented as 6 dimension sets (see Table 2). Results shown in the table are standardized in order to facilitate visualization and tool development in the later section. Through investigation into those numbers alone, a teacher can provide specific advice to a particular group. For example, if the value of a student in the *Community* dimension (Group 4) is very low, the teacher could suggest that group communicate more among team members. Similarly, if the group has a high value on Subject (Group 6), then it indicates that group members are independently active.

Table 2. Activity theory informed sample group measurements

Dimension / **Group**	Subject	Rules	Tools	Division of Labor	Community	Object
Group 1	-0.6293	-0.4625	0.1535	1.4359	1.3411	-0.3078
Group 2	-0.6667	-0.5365	0.5637	0.3860	-0.3055	0.4673
Group 3	-0.2755	0.1475	0.3643	-0.5009	-0.3404	-0.3780
Group 4	-0.5633	-1.0875	-0.9616	-0.3951	-1.0344	-1.5361
Group 5	1.4729	1.2876	-0.8774	1.3323	-0.9299	1.8201
Group 6	2.7025	1.4486	-0.8774	-1.0588	-0.3491	1.7011
⋮	⋮	⋮		⋮	⋮	⋮

Cluster Results

Using the K-means algorithm, an arbitrary number of clusters are formed for further analysis of results. Table 3 shows final clustering results.

Table 3. Cluster results

Group Clustering Summary		Cluster1		Cluster2		Cluster3	
		2/14		4/14		8/14	
Measure	Range (Standardized)	Mean	SD	Mean	SD	Mean	SD
Subject	[-1.343, 2.702]	2.012	0.838	-0.637	0.560	-0.185	0.442
Rules	[-1.300, 2.102]	-0.845	0.000	1.065	1.018	-0.321	0.666
Tools	[-1.034, 2.160]	-0.616	0.396	1.367	0.731	-0.529	0.330
Division of	[-2.079, 1.449]	1.323	0.110	-0.317	1.227	-0.172	0.787

Labor							
Community	[-1.130, 2.332]	0.614	2.311	-0.225	1.129	-0.041	0.640
Object	[-1.536, 1.902]	1.745	0.149	-0.217	0.560	-0.328	0.837

Cluster 1: Individually Participative and Great Community but Less Adept in the Environment.
This cluster contains 2 groups, in which members of these groups were actively involved in the modules as individuals, because they had the highest value (2.012) in the *Subject* dimension. Similarly, these two groups also ranked the highest in *Division of Labor* (1.323) and *Community* (0.614), reflecting that they shared the workload fairly equally and maintain community structure well. Furthermore, they also received the highest value on *Object*, indicating that these two groups were very active across all modules. By contrast, these two groups received the lowest score on *Rules* (-0.845) and *Tools* (-0.616) respectively. Teachers could infer that students in these groups might heavily rely on one or two functions or tools that VMT provides. Therefore, they can encourage students to further explore the VMT environment.

Cluster 2: Great Explorer into the Environment but not Great Group Learners.
This cluster has 4 groups in it and seems to have almost the exact opposite results as compared to Cluster 1. They have the best performance in *Rules* and *Tools* measurements, indicating that students in these groups used a variety of tools or functions VMT offers. However, from individual and group learning perspectives, these groups did not perform well enough. In the *Subject* dimension, they received the lowest value (-0.637) showing that students invested significant personal effort; in the group dimension, students ranked the lowest both on *Community* (-0.225) and *Division of Labor* (-0.317) demonstrating that students in these groups did not make equal contributions to problem solving and did not communicate well with group members. This cluster of groups ranked middle in the *Object* dimension indicating that these groups were relatively active across all modules.

Cluster 3: Average Performance as a whole but not Consistently Active
Most of the groups (8 of 14) fell into this cluster, which has average performance. *Subject* (-0.185), *Rules* (-0.321), *Tools* (-0.529), *Division of Labor* (-0.172), and *Community* (-0.041) all ranked in the middle overall groups. In contrast with these dimensions, *Object* has the lowest value (-0.328). This showed that students in those groups are probably very active in some of the modules but limited participation in others. This is especially true when considered together with the average performance across all other 5 dimensions. While we expect this *Object* would rank in the middle as well, it countered our expectation. Therefore, teachers may want to discover which module these groups were not active in and check their curriculum design.

Validation and Comparison with Qualitative Chat Analysis
To validate our quantitative methodology, qualitative analysis of chat data from Topic 5 was implemented. Topic 5 is mainly designed for students to explore the dependences of geometric objects and build an equilateral triangle inscribed in another equilateral triangle. Using the method described above, these groups have first been put into three different clusters solely based on their collaboration on Topic 5. Then three groups were chosen, one per cluster, and compared with the qualitative analysis results of the three groups. Qualitative chat analysis from an ethnomethodological perspective has been established for years in the context of VMT (Stahl,

2006, 2012), which is a sociological approach to describe the work that people do in collaboration with others to build social order and meaning (Garfinkel, 1967). In contrast with conversation analysis with a focus on construction of "adjacency pairs," short sequences in which one person's utterance ignites a response (e.g. an instance of question-answer pair), ethomethodological chat analysis look for longer sequences, in our context a series of math proposal adjacency pairs (Stahl, 2006, Ch. 21 esp. pp. 442-456), of examples of mathematical problem solving by groups (Stahl, 2012). For our chat analysis, we used the VMT Replayer which allows us to see chat and recorded actions at the same time (Figure 2).

Group 14 contains four members (s50, s51, s52, s53) and conversation began with a proposal from a member "Okay lets start by taking turns dragging vertex A and vertex D." Students in this group took turns proposing actions and building geometric objects, and in the end successfully complete the module objective and are therefore considered a high performing group. This result aligns well with our activity-theory based assessment, which assigns group 6 to the high cluster and the values of 6 dimensions of activity theory are all above class averages. The group as a whole collaborated very well when helping one another use the circle tool correctly in order to make sure the segment they created was the radius of both circles (Figure 6). Specifically, one group member (s50) invested great effort in explaining the process step-by-step for group member (s51) to follow, but she did not succeed. Then s50 helps her teammate, making the circles and intersecting points.

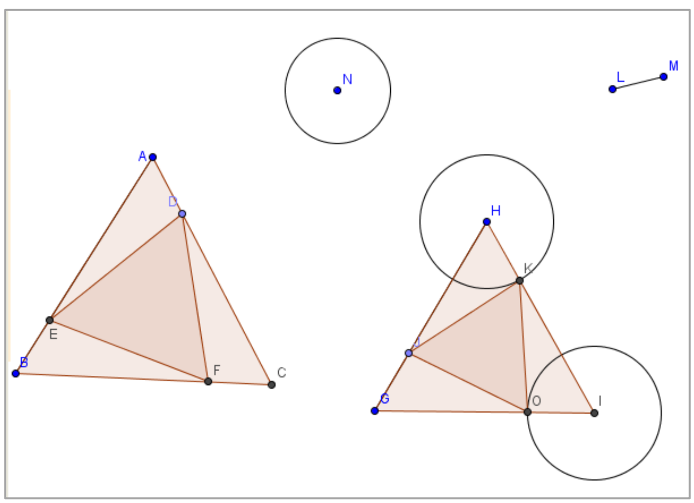


Figure 6. Group 14

Group 4 has three members (s13, s14, s15) and began by establishing a common ground, agreeing to take turns exploring dependencies in the environment. This group is in the "medium" performing cluster, and although two of the three members came close to completing the assignment correctly, their answer missed key dimensions. Coordination between two of the members went smoothly on moving the vertex of the triangle, but group member (s13) was reluctant in relinquishing control to other members with "Fine" when it is time for others to try. They used correct terms like "equidistant" and "construction" to communicate. Then one of the members proposed to start construction of the two triangles. At the beginning group member (s13) made all the constructions of the triangle without explaining to the other members though other members (s14) asked him to do so “kk tell us what ur doing”. Then group member (s15) asked for control and followed an example taught in class using circles (Figure 7).

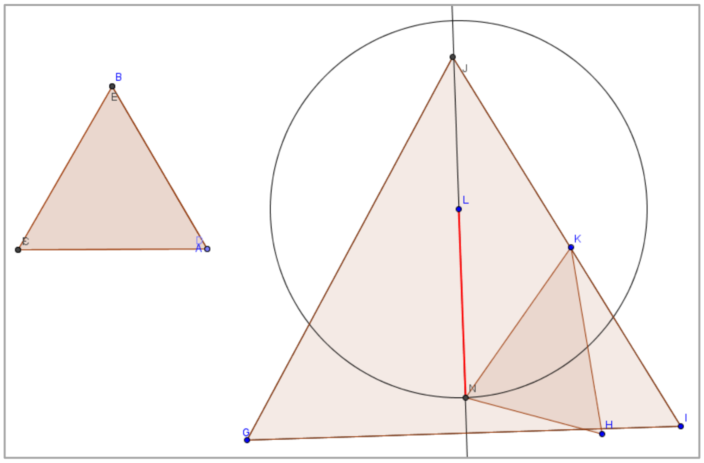


Figure 7. Group 14

Group 12 consisted of four members (s40, s41, s42, s43) and started by moving the vertex, discussing the dependencies they noticed for each triangle. They used many geometric terms but did not typically “notice and wonder,” or follow through on their ideas. This group is from the low performing cluster, and both failed to successfully complete the task, and exhibited many difficulties in collaboration. For example, this group did not discuss why they needed to use the Circle function to make the equilateral triangle. One team member (s40) commented “so I think triangle igh is like triangle abc” (Figure 8), again, the team did not discuss why this is so and no member responded to this comment. After they completed the outer triangle, they attempted to build the inner triangle. They talked about how f and e are dependent on d (Figure 8).

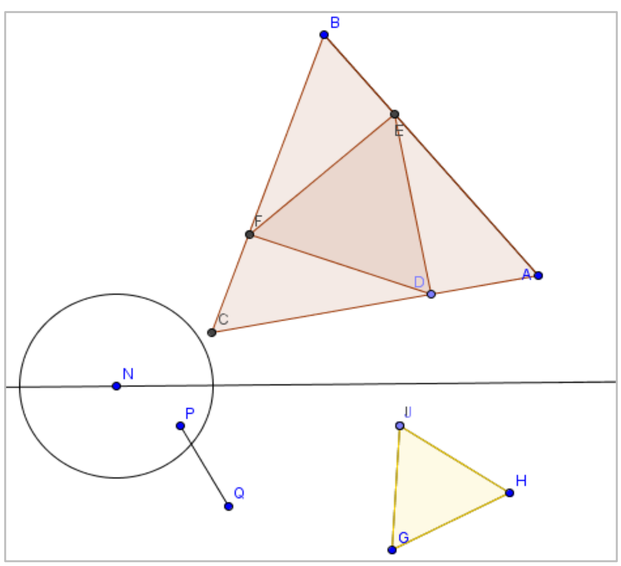


Figure 8. Group 12

In sum, our quantitative analysis obtained similar conclusions with the qualitative chat analysis. Since the proposed method can be totally automated, compared with qualitative method, it is much more efficient and practical in supporting teachers' decision making because our approach is able to support real time assessment and enable teachers to provide immediate feedback.

Tool Development

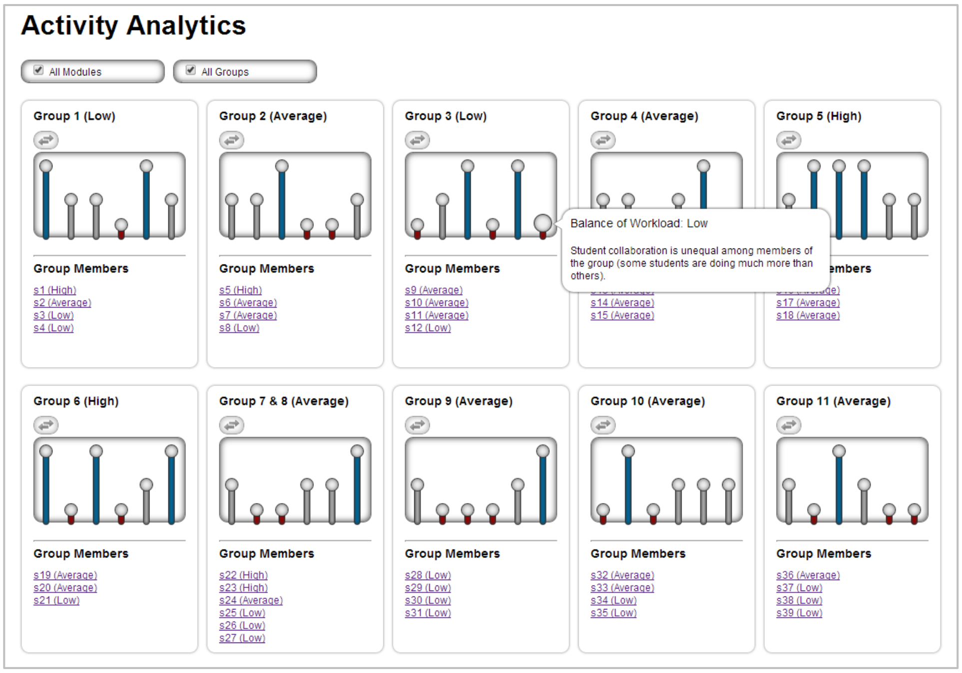
It would be difficult in practice for teachers to infer meaning or make sense of figure 12 without a technical or analytical background. Instructors can more easily obtain information on student participation from these results using visualizations. This is especially true when the number of students is large. The research community in learning analytics has recognized the significance of visualization, establishing it as a required component in the learning analytics cycle (Shum & Ferguson, 2012). In addition, a graphical display allows teachers to process information more effectively than do pure numbers (Leinhardt et al., 1990, Goggins, Xing, Chen, Chen, Wadholm, in press).

On the other hand, teachers cannot identify specific problems simply based on the clusters of groups. For example, in Cluster 2, teachers may conclude that some students are not as active as others. How can the teacher specifically identify who these less active students are? In Cluster 3, teachers may infer that students are participating in some of the modules but not others, but does not know which modules the students are involved in. A visualization tool would be helpful in this situation. In combination with our previous work on individual assessment (Xing, Wadholm, & Goggins, 2014a & 2014b), we used graphs to present the quantified 6 dimensional data for both individuals and groups. This allows teachers to see a particular cluster of groups across any module.

Table 4. Tool functionality

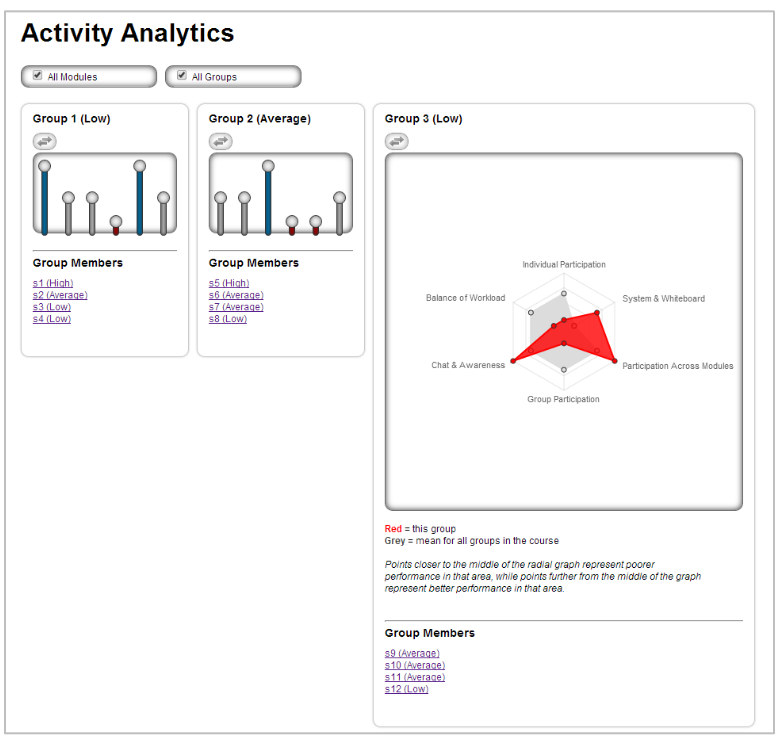
Data / **Type Unit**	Activity Theory	Cluster	Module	Groups
Individual	6-dimensional data	3 clusters	Any Module or Combination of Modules	Any Group or Combination of Groups
Group	6-dimensional data	3 clusters	Any Module or Combination of Modules	Any Group or Combination of Groups

Figure 9 shows the general interface and basic functions of an early draft of the web application, which allows teachers to access this detailed information from anywhere and at any time. With the group, individual, and module information together, the teacher can identify issues with individuals, groups, and the class as a whole in order to provide concrete advice and interventions.



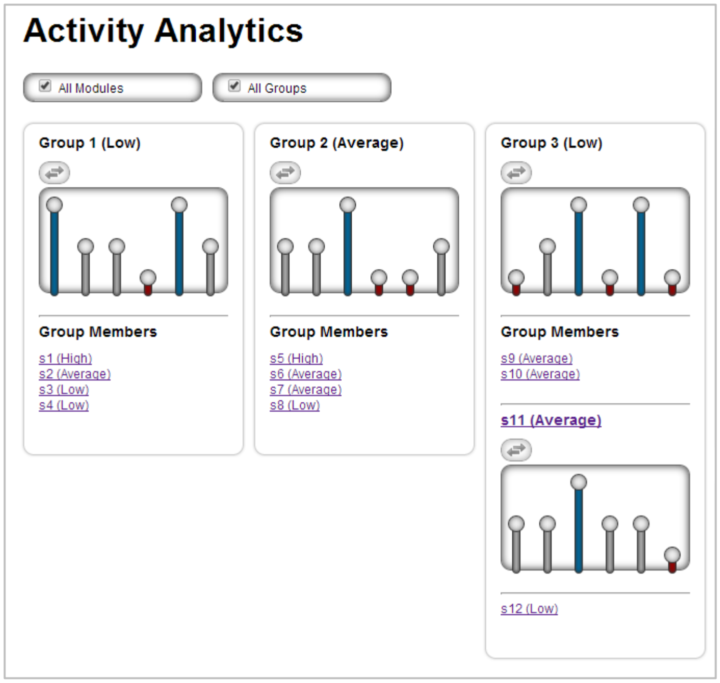
(a)

Figure 9a: information on participation across all groups and individuals in 5 modules. It gives an overview of which clusters these groups belong to (High, Average, or Low), the members of these groups, and the clusters each student is individually affiliated with. The six bars with nodes at the top demonstrate information on details of the six dimensions of activity theory for a group. The teacher is able to see how the group performs in each dimension by hovering over nodes in the graph. Since the naming of the six dimensions of activity theory (*Subject, Tools, Object, Rules, Community, Division of Labor*) is relatively abstract, we replaced the six dimensions with Individual Participation, White Board & System, Participation across Modules, Group Participation, Chat & Awareness, and Balance of Workload, for teachers to better understand the results. Hovering over a node shows more detailed information.



(b)

Figure 9b: Also, there is a button for the teacher to switch visualizations in order to compare the performance of each dimension with the mean for the course.

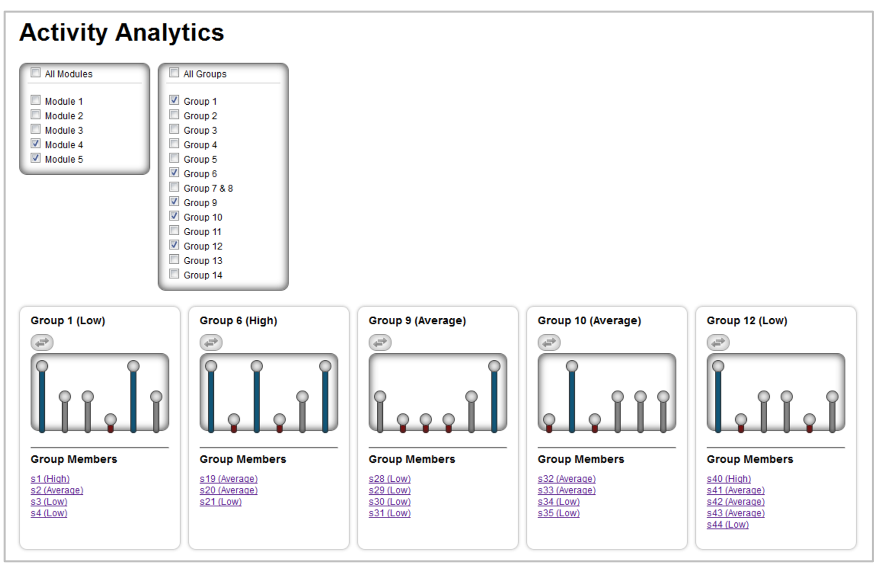


(c)

Figure 9c: After seeing the big picture of how each group performs, teachers could obtain more detailed information on how each individual in the group performs. Individual overall performance is identified as being a part of the High, Average, or Low cluster, and further information is shown by clicking on the name of the student

(d)

Figure 9d: Teachers can click the Switch button to see the radial graph and compare the student's performance with the mean for the course.



(e)

Figure 9e: On the other hand, teachers could also decide to view the performance of groups and individuals over a particular module, or between specific groups

Figure 9. VMT Activity Analytics Tool

Discussion & Conclusion

In this paper we explore a different way of developing assessment that is scalable and coherent with theory. Activity theory holistically frames group participation in our framework, which addresses situatedness, contextuality, and social mediation. In addition, activity theory looks at ongoing processes to explain group changes and performance over time (De Laat, 2006). This is important especially considering that a major portion of group assessment in CSCL is still summative in nature (e.g. tests, product evaluation). On the other hand, researchers have created a variety of constructs to assess collaborative processes to infer group learning. These measurements are usually derived from analysis of chat and/or tool usage and therefore unable to assess group learning fully. 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 teacher using this method to provide timely feedback to individuals and groups. In addition, these coding methods and frameworks are not always shared (Gress et al, 2010) which leads to difficulty in maintaining consistency across different evaluators. The proposed method, from measure construction to clustering and visualization, is totally automated. Therefore, it could significantly reduce the teachers' assessment burden and offer timely feedback based on assessment results. Automatic methods could also increase the consistency of evaluations and improve the reliability of results.

Though our method is also computational in nature, it is informed by activity theory, which centers on descriptions of human socio-technical behavior. Unlike previous quantitative methods that focus on one aspect of participation such as number of posts, login times or clicks on a page (Wolff, Zdrahal, Nikolov, & Pantucek, 2013), this theory provides a holistic view of a student's participation in CSCL activities. While a purely computational method makes it difficult for a teacher to offer tangible advice to students to improve their performance, activity theory provides a semantic background for the teacher to give specific and meaningful feedback to students. Moreover, many times, computational results are difficult for teachers to use, comprehend and explain (Ferguson, 2012; Xing, Guo, Petakovi & Goggins, 2014). To facilitate teachers' timely interpretations of results, we developed a web-based tool that visualizes the data. Therefore, we are exploring natural language processing of the chat log data and incorporating it into our activity theory measure construction system and visualizations in order to further inform learning 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).

Calvani, A., Fini, A., Molino, M., & Ranieri, M. (2010). Visualizing and monitoring effective interactions in online collaborative groups. *British Journal of Educational Technology*, *41*(2), 213-226.

Daradoumis, T., Martínez-Monés, A., & Xhafa, F. (2006). A layered framework for evaluating on-line collaborative learning interactions. *International Journal of Human-Computer Studies*, *64*(7), 622-635.

De Laat, M.F. (2006). Networked Learning. Apeldoorn: Politie Academie. *PhD dissertation*, Utrecht Univ., the Netherlands.

Engeström, Y. (1987). Learning by expanding: an activity-theoretical approach to developmental research. Helsinki: Orienta-Konsultit Oy.

Fessakis, G., Dimitracopoulou, A., & Palaiodimos, A. (2013). Graphical Interaction Analysis Impact on Groups Collaborating through Blogs. *Educational Technology & Society*, *16*(1).

Ferguson, R. (2012). The state of learning analytics in 2012: A review and future challenges. *Knowledge Media Institute, Technical Report KMI-2012-01*.

Frederiksen, N. (1984). The real test bias: Influences of testing on teaching and learning. *American Psychologist*, *39*(3), 193.

Goggins, S., Xing, W., Chen, X., Chen, B., & Wadholm, B. (in press) Learning Analytics at "Small" Scale: Exploring A Complexity-Grounded Model for Assessment Automation. *Journal of Universal Computer Science*.

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.

Guo, R., & Zhang, Y. (2014). Identifying Time-of-Day Breakpoints Based on Nonintrusive Data Collection Platforms. *Journal of Intelligent Transportation Systems*, *18*(2), 164-174.

Kapur, M., & Kinzer, C. K. (2009). Productive failure in CSCL groups. *International Journal of Computer-Supported Collaborative Learning*, *4*(1), 21-46.

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.

Leinhardt, G., Zaslavsky, O., & Stein, M. K. (1990). Functions, graphs, and graphing: Tasks, learning, and teaching. *Review of educational research*, 60(1), 1-64.

Lipponen, L., Rahikainen, M., Lallimo, J., & Hakkarainen, K. (2003). Patterns of participation and discourse in elementary students' computer-supported collaborative learning. *Learning and instruction*, *13*(5), 487-509.

Milrad, M. (2002). Using construction kits, modeling tools and system dynamics simulations to support collaborative discovery learning. *Educational Technology & Society*, *5*(4), 76-87.

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.

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.

Shum, S. B., & Ferguson, R. (2012). Social Learning Analytics. *Educational Technology & Society*, *15*(3).

Stahl, G. (2006). *Group cognition: Computer support for building collaborative knowledge*. Cambridge, MA: MIT Press.

Stahl, G. (2012). Theories of cognition in collaborative learning. *International handbook of collaborative learning*, 74-90.

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

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard university press.

Wolff, A., Zdrahal, Z., Nikolov, A., & Pantucek, M. (2013). Improving retention: predicting at-risk students by analysing clicking behaviour in a virtual learning environment. In *Proceedings of the Third International Conference on Learning Analytics and Knowledge* (pp. 145-149). ACM.

Xing, W., Guo, R., Fitzgerald, G., & Xu, C. (2014). Google Analytics based Temporal-Geospatial Analysis for Web Management: A Case Study of a K-12 Online Resource Website. *International Journal of Information Science and Management (IJISM)*, *13*(1).

Xing, W., Guo, R., Lowrance, N., & Kochtanek, T. (2014). Decision Support Based on Time-Series Analytics: A Cluster Methodology. In *Human Interface and the Management of Information. Information and Knowledge in Applications and Services* (pp. 217-225). Springer International Publishing.

Xing, W., Guo, R., Petakovic, E., & Goggins, S. (in press). Participation-based student final performance prediction model through interpretable Genetic Programming: Integrating learning analytics, educational data mining and theory. *Computers in Human Behavior*.

Xing, W., Wadholm, B., & Goggins, S. (2014a). Learning analytics in CSCL with a focus on assessment: an exploratory study of activity theory-informed cluster analysis. In *Proceedins of the Fourth International Conference on Learning Analytics And Knowledge* (pp. 59-67). ACM.

Xing, W. L. Wadholm, B. & Goggins, S. (2014b). Assessment analytics in CSCL: Activity theory based method. In *Proceedings of the international conference on learning sciences'14* (pp. 1535–1536), Boulder, Colorado, USA.