Learning Analytics in CSCL with a Focus on Assessment: An Exploratory Study of Activity Theory-Informed Cluster Analysis

Wanli Xing
University of Missouri-Columbia
School of Information Science & Learning Technologies
1(281) 309-8515
wxdg5@mail.missouri.edu

Bob Wadholm
University of Missouri-Columbia
School of Information Science & Learning Technologies
1(812) 391-2017
rwadholm@indiana.edu

Sean Goggins
University of Missouri-Columbia
School of Information Science & Learning Technologies
1(573) 882-7807
gogginss@missouri.edu

ABSTRACT
In this paper we propose an automated strategy to assess participation in a multi-mode math discourse environment called Virtual Math Teams with Geogrebra (VMTwG). A holistic participation clustering algorithm is applied through the lens of activity theory. Our activity theory-informed algorithm is a step toward accelerating heuristic approaches to assessing collaborative work in synchronous technology mediated environments like VMTwG. Our Exploratory findings provide an example of a novel, time-efficient, valid, and reliable participatory learning assessment tool for teachers in computer mediated learning environments. Scaling online learning with a combination of computation and theory is the overall goal of the work this paper is situated within.

Categories and Subject Descriptors
J.1 [**Data Processing**]: Education; K.3.1 [**Computers and Education**]

General Terms
Algorithms, Measurement, Design, Reliability, Assessment, Human Factors, Theory.

Keywords
Learning Analytics; CSCL; Activity Theory; Educational Assessment

1. INTRODUCTION
"Learning Analytics is the measurement, collection, analysis and reporting of data about learners and their contexts, for purposes of understanding and optimizing learning and the environments in which it occurs" [1]. While measurement and assessment of learning is a major goal for learning analytics, it is also a demanding experience for many teachers. Assessment is not just important for evaluating learning outcomes: it may also be motivating for many students who have a performance goal orientation [2]. Assessment of learning in computer-supported collaborative learning (CSCL) environments is more than merely measurement of outcomes; the quality of the collaborative learning processes [3] is also salient.

While the diverse theoretical and methodological positions in CSCL are well studied [4, 5, 6], common assessment practices in CSCL remain largely unexplored [3]. When looking at assessment methods, teachers need to consider practical issues such as time, validity, reliability, and individual accountability. Teachers are under time constraints, so assessment methods should allow them to evaluate learning and participation in CSCL activities in a timely manner [2, 7]. Reliability is "an indication of the consistency of scores across evaluators or over time" [8]. For an assessment to be reliable, significantly similar results must occur each time the assessment is performed, regardless of who is involved in the assessment or when the evaluation occurs. Validity is how well an assessment actually measures what it is supposed to measure [8]. In a CSCL context, validity means whether the measures actually assess learning. Individual accountability reflects whether the assessment influences individual performance [9, 10]. Current assessment strategies employed by teachers in CSCL have difficulty meeting these four requirements.

After a review of 186 articles and 340 measures incorporated in CSCL, Gress et al. [11] grouped assessment into seven categories: self-report, interview, observation, process data, discussions and dialogues, performance and products, and feedback and/or prediction. These assessment methods commonly in use today tend to violate the assessment requirements of time (several of these are very time intensive), validity, reliability, and individual accountability in one way or another. For instance, self-reports (usually in the form of reflections, surveys, or questionnaires) might be time efficient, but might also be too subjective and affect the validity and reliability of an assessment [12].

Interview, observation and coding of process-oriented data (discussions and dialogues) are usually time-intensive and add a huge burden to the already heavy duties of teachers. Additionally, due to the lack of a standard metric or published set of descriptions, these methods may lead to difficulties in determining the validity of those assessment methods and therefore raises questions in terms of identifying and defining the quality of learning and performance [11]. Many constructs, protocols, and schemes reflect a lack of replication and examination of reliability across teachers and contexts [11]. Specifically, with discussions, dialogue or content analysis, most assessment practices do not directly measure students' interactions at face value as technological manipulation or in relation to knowledge construction.

LAK'14, March 24–28, 2014, Indianapolis, IN, USA

http://dx.doi.org/10.1145/2567574.2567587

Teachers' engagement in performance and product assessment usually consists of individual task evaluation (typically tests, quizzes, essays etc.), and group evaluation (all output produced by students' collaborative activities), with each student receiving the same grade or individual grades based on a combination of group and individual tasks. On one hand, this way of assessment may suffer from significant flaws in assessing students' learning. Tests may not be the best indicator of learning. These methods may overlook the process and context information that students might gain during the CSCL process; one of the supporting philosophies in CSCL [2].

On the other hand, grading of group work can be problematic, both practically and theoretically [2, 11]. Several issues that have been identified with group grading include: 1) individual accountability is lost and freeriding often occurs (at the expense of one or two in the group who do most or all of the work) [13, 14], 2) group grades may tend to over- or underspecify an individual students' competence (equal work may receive different grades only because of the makeup of a student's group, which may be outside of their control), 3) lower performing students may benefit more from group grades than their high performing counterparts [15], and 4) unsatisfactory experiences with group grades may end up causing students to dislike collaborative learning activities [3].

In daily practice, the most popular assessment method in the feedback category is peer assessment. In this approach, focus is on peer ranking (peer work in different areas is ranked on scale), peer nomination (highest performing members of a group are nominated), or peer rating (each group member rates all other members) [7]. Nominations, rankings and ratings have all been found to result in strong adverse reactions by students [16, 17]. Pond, Ul-Haq and Wade [18] argue that these methods are all prone to bias, distinguishing four origins for bias: over-marking ('friendship marking'), purposive lack of differentiation within groups (collusive marking), dominant individuals receiving highest marks (decibel marking) and non-contributing students benefiting from group marks (parasite marking). Even though peer assessments save teachers time, the reliability and individual accountability in these methods may be in doubt.

Assessments of learning in CSCL from process data usually takes the form of estimated task times, frequency of participation, and sequencing of events, in addition to trace data [11, 19]. This method is time efficient, reliable and accounts for each individual. Unfortunately, many of these assessments end up with mere participation measures (time on task, messages posted), which may or may not be good indicators of learning [2]. Therefore, there may be serious validity issues with these forms of assessment. Other assessment methods such as navigation path or social network analysis (content analysis) [20] suffer from some of the same problems.

Even though Gress [11] identified seven categories of assessment methods, teachers tend to rely on conventional text-based measures like self-reports, observations, and content analysis rather than using automated learning analytics to assess specific details of learning. Those traditional methods are not only time consuming but may also suffer from insufficient validity, reliability, and individual accountability as outlined above. Something more is needed. The paucity of the use of original methodologies and research done in this area is not due to irrelevance or lack of interest: rather, it seems that teachers are stymied by a lack of proven methodologies and paradigms that can be applied in a straightforward and holistic way.

However, in an educational research context, there are indicators that all roads lead back to learning [2]: even though current assessment methods are not perfect, their measures and frameworks provide insights for teachers to understand students' learning. As a result, the issue of learning assessment in CSCL becomes that of building on a proven theoretic paradigm, using holistic measures (incorporating various indicators and measures used by previous methods), and providing a time-efficient, valid, and reliable learning assessment tool to teachers that ensures individual accountability. Further, Dennen [2] argues that assessment requires "consideration of the learning activity and medium, as well as the resulting artifacts."

This exploratory study presents a unique application of activity theory applied to the assessment of CSCL. It is meant to offer a methodology that can help to fill in the gap of assessment of student participation in learning activities in CSCL environments, by assessing student activities holistically and by using cluster analysis to evaluate strengths and weaknesses in individual students' participation in collaborative activities.

This paper is organized as follows. First, we will provide a background on activity theory and describe how it is a useful framework in this context. Next, we will directly apply activity theory to a CSCL environment, and provide a specific case for application of this assessment methodology, analyzing and interpreting holistic participation of individual students in CSCL activities. Last, we will explicate strengths and weaknesses of such an approach, and provide directions for future research.

2. THEORETICAL FRAMEWORK

2.1 Activity Theory

Activity theory is a psychological and multidisciplinary theory that seeks to be naturalistic and offers a holistic framework for describing activities in practice while linking together individual and social behavior [21, 22, 23, 24]. A model of the structure of an activity system was formulated by [22], and includes the interacting components of subject, object, tools, division of labor, community, rules, and outcome (see Figure 1).

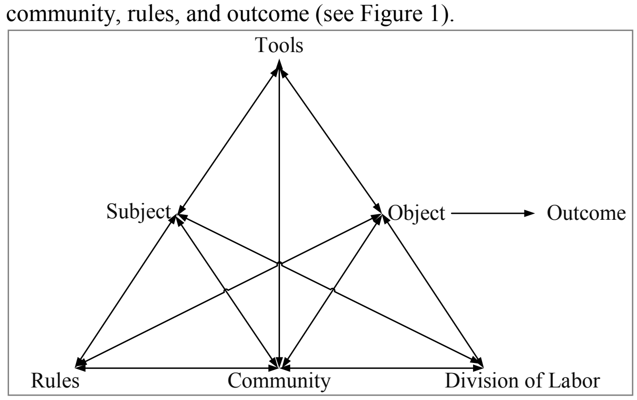

Figure 1. Activity Theory

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" [25]. Activity theory focuses on how participants transform objects and how components in a system mediate this transformation [22]. Learning is reframed as social practice (rather than as merely the product of practice) [22, 26]. In our CSCL learning assessment context, the outcome and process of this transformation are learning and performance. It is the sum of the system components and the tensions among them that make up the learning and influence the learning outcomes. Current common assessment methods only consider part of the activity of the learning system.

In addition to the interactions of components within an activity system, a system can have subsystems, which can be separate (full-fledged) activity systems or other instances of the same system depending on perspective [23]. Therefore, in order to highlight the learning outcome of an individual student to facilitate assessment in individual accountability, the activity system can be thought of as being built for each student. Thus, the activity of learning in a CSCL context for every student can be presented as:

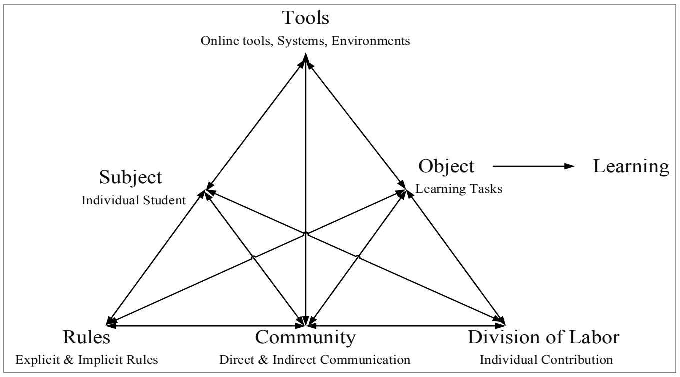

Figure 2. Activity Theory in CSCL assessment

Table 1. Measure-metrics from Activity Theory

Measure-metric	Definition
Object	Solve learning tasks such as solving a problem or producing an artifact (e.g. essays)
Subject	Individual student involved in this activity. When assessing learning, an individual's differences of effort, motivation, roles etc. should be taken into account
Tools	Computers, online tools, systems, and environments that mediate the learning activity
Community	Direct and indirect communication enables an individual subject to help maintain a sense of community with other students, teachers, and support staff
Rules	Implicit and explicit rules and guidelines that constrain the activity. For example, teachers can set specific rules for a learning task (explicit) and an individual student can only use the functions residing in the supporting tools (implicit)
Division of Labor	Concrete contribution each individual makes to the overall object

2.2 Case Description

To operationalize Activity Theory as a lens for making sense of electronic trace data from a Synchronous Math discussion board, we focused on several modules of a course designed to be taught with Virtual Math Teams with Geogebra (VMTwG) software. The five modules we analyzed included teams of three to four practicing teachers who were learning how to implement the curriculum in VMTwG while going through the curriculum themselves. The five modules we analyzed include: a warm up activity, "Messing around with dynamic geometry", "Visualizing the world's oldest theorems", "Constructing triangles" and "programming custom tools". The full curriculum currently includes a total of 18 topics, and is available at the project website (http://vmt.mathforum.org).

3. METHOD

In this section, we reference activity theory to model individual student performances in CSCL activities.

3.1 Dataset Description

We collected all the log data for this study in .txt format, which centers on specific event types from the CSCL environment (VMT): Awareness, Geogebra, System, Chat, 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, or update a segment etc.). The System event type records information on how the VMT environment is accessed. For example, a student joins a room, leaves a room or views different tabs created by the students or teachers. Wb logs actions on how tools are being used in the white board areas, such as resizing of objects, creating a textbox, etc.

For every event type, we have logs of what action (adding a point, sending a chat, erasing a message, or creating a text box, etc.) the student makes under what subjects/tasks (modules and tasks). In addition, the environment logs the information about when this action takes place (time) and how long it lasts (duration) as well as in which virtual room the event occurs.

3.2 Activity Theory and Behavioral Data

Establishing theoretical coherence between the behavioral data analyzed and the theoretical lens is a significant concern for socio-technical systems research [27]. As Howison et al note [27], a good deal of prior research is not explicit regarding this connection, which has the deleterious effects of drawing findings into question. In the analysis of group based interaction data, the Group Informatics ontology and methodological approach provides researchers with a set of patterns to follow when drawing connections between behavioral trace data and a particular theory [28].

In this study we relied on the third authors five years of participation in the development and analysis of small group learning in VMTwG to make empirically informed choices about how to map trace data from the system into an activity theoretic analysis framework. The exploratory work reported here is a step forward in the development of instrumentation (learning analytics) for small group learning systems like VMTwG. We analyzed our results reflexively in the context of ongoing research and development in VMTwG.

3.3 Dataset Preprocessing

3.3.1 Remove Noise Data

Because various teachers and VMT team members may join the class from time to time to either facilitate the students' learning or test the environment, our first step is to remove outsiders' data. Next, VMT has different topics and modules. Some tasks or modules are not related to learning (such as the Focus Group Interview), and some are topics not covered in the class (such as Trigonometry). We deleted the log data generated around the irrelevant modules.

3.3.2 Data Categorization

Since the assessment of learning is more summative in nature, we did not consider the time dimension. However, our method enables teachers to assess students' learning at any time during the course to better monitor students' participation in CSCL activities. For the purposes of this study, we only considered the situation at the end of the training to assess students' participation in collaborative learning activities.

As stated above, assessment should be designed for each student to meet the individual accountability requirement. Therefore, all the data are prepared for an individual student. Since the log data is centered around event types, as well as to facilitate measure construction in the next section, we process each event into four dimensions (individual, group, action constraints, and module set) denoted as below:

E_{1j} =Awareness = {individual, group, action constraints, module set}
E_{2j} =Chat = {individual, group, action constraints, module set}
E_{3j} =Geogebra = {individual, group, action constraints, module set}
E_{4j} =System = {individual, group, action constraints, module set}
E_{5j} =Wb = {individual, group, action constraints, module set}

Some modules require personal exploration or experimentation in the VMT environment and some modules ask the student to collaborate with others to solve a problem or produce an artifact. Therefore, the individual category sums all the personal endeavors or actions (frequency in each event type). Similarly, the group category sums all the actions the student makes in the group project (frequency in each event type). The action constraints dimension sums all the types of action the student performs under that event type over the whole training. For example, if a student never erased a message in the Awareness event over all the modules, then the action constraints for Awareness is 0; for a Wb event, if a student takes actions such as creating a textbox, or copying an object, but never uses other actions such as moving objects, resizing, etc. across the whole class, then the action types equals 2. Some students may miss one or two modules. Therefore, the module set dimension records what modules the student is involved in under each event type. Rather than a single value, each module set is comprised of the modules in which events take place. E denotes the event type, E_{ij} the value of every dimension in every event type, where *i* denotes the event type $i \in [1, 5]$ and *j* denotes the dimensions, $j \in [1, 4]$, $i \in Z^+$, $j \in Z^+$.

Therefore, an event type *i* can be presented as: E_i = *{individual, group, action constraints, module types}* $= [E_{i1}, E_{i2}, E_{i3}, E_{i4}]$

3.4 Measure Construction

Subject: In the CSCL activity of learning (Figure 2), Subject represents the individual student. When mapped to our log data, it represents all actions one student makes during the whole training under the individual modules and tasks, detonated as E_{i1} . In other words, it is the sum of all actions under individual tasks. Therefore, for every student n, Subject can be denoted as: Subject $X_{n1} = \sum_{i=1}^{5} E_{i1}$ (assuming there are N students in total, n*=1, 2, ⋯, N*).

Rules: As shown in Figure 2, Rules includes implicit and explicit rules. Under the social-technical construct, the rules are the implicit rules that constrain students' actions, denoted as E_{i3} . In this VMT context, students have to perform actions that the VMT environment offers. Therefore, the rules are reflected by the actions the student uses across all the modules as: Rules $X_{n2} = \sum_{i=1}^{5} E_{i3}$.

Tools: In VMT, Tools are tools that facilitate the learning activity. Under the VMT context, the tools are the System and Wb where the student's action for tool usage is registered, denoted as E_{i1} and E_{i2} . Tools $X_{n3} = \sum_{i=4}^{5} (E_{i1} + E_{i2})$.

Community: All the communications that help maintain the community structure. In terms of the VMT context, students 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 E_{i1} and E_{i2} . Therefore, Community can be presented as the sum of chat messages and awareness. Community $X_{n4} = \sum_{i=1}^{2} (E_{i1} + E_{i2})$.

Division of labor: Division of Labor is the contribution of the student made to the collaborative learning modules. Even though chat messages may also contribute to the geometry object development, the concrete contribution to the geometry object construction is from the Geogebra dimension. As a result, the division of labor can be denoted as Division of Labor $X_{n5} = E_{32}$.

Object: The CSCL activity is to achieve the object of a student's active involvement in the whole class. Hence, the first factor to consider is the number of modules students participate in. In order to quantify whether the student is active in those learning modules, we incorporate the totally frequency of participation and the number of event types. By doing this, we can avoid too high of ratings for the student who participates in all the modules but makes very few actions or contributions.

On the other hand, the number of modules the student is involved in is the key for the object dimension; the other two are secondary factors. In addition, because the number of modules is in the scale of 10, while the frequency for participation is in 1000, we use the *log* function to dampen the effect of the frequency measure. Even though the event types are in the same scale of modules, we still want to dial down a little for overall event types that the student used. Thus, we use a fraction to lower the event type effect for object measurement, characterized as the event type students are involved in divided by the overall event types (5). In sum, we denoted Object as

Object $X_{n6} = \log \sum_{i=1}^{2} (E_{i1} + E_{i2}) * \left| E_{14} \cup E_{24} \cup E_{34} \cup E_{44} \cup E_{54} \right| * (\left| E_i \right| / 5)$

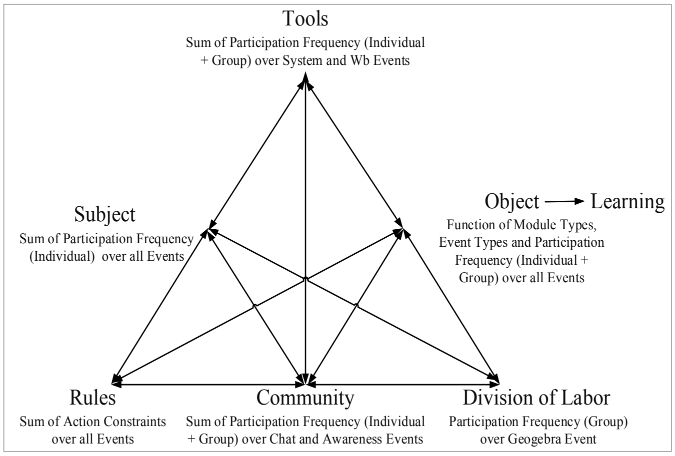

Figure 3. Activity Theory in VMT Assessment

In sum, based on activity theory, we have built a quantified model for individual student performance in CSCL activities specific to the VMT environment: [Subject, Rules, Tools, Community, Division of Labor, Object], as illustrated is Figure 3.

3.5 Assessment of Participation in Collaborative Learning Activities

3.5.1 Rationale

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 learning. Therefore, it is hard to compute one value as functions of the six dimensions to indicate the learning or performance result of an individual student, especially considering the complexity of the nature of learning [2, 32]. For example, even though a specific student contributes much in the Division of Labor (has a high value in that dimension), we cannot necessarily assume this student may attain learning from this activity. This In student might participate in only one learning module, which makes him or her deficient in the Object dimension. In other words, this student does not get exposed to all the knowledge (various learning modules) but just actively participates in one particular module.

In this exploratory study, we used K-means cluster analysis to group students with similar learning results. A K-means cluster brings into consideration all 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. It is hard to make any inferences based on an individual student's performance. Clustering can help the teacher to see the general pattern of students' performance levels and grasp the big picture of how students are participating in collaborative work in the course.

3.5.2 Procedure

Cluster analysis, which addresses the problem of data segmentation, belongs to unsupervised learning methods since there is no knowledge of "preferred" clusters [29]. 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 [30]. A significant step in clustering is to define the system scale and select the proper cluster elements, which have been defined in the last step-measure construction. By considering the measurements defined in the subsection of measure construction, it is possible to group the students into different categories with a mathematical method. Therefore, the state definition used for this study is a vector of measurements for all the students. The data samples are therefore multidimensional because the vectors of measure for each student 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 there are *N* students. For instance, the constructed measure can be recorded for N individuals (students). Then the data **X**, a *K (6)* × *K (N)* matrix, will have the format as following.

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

To deal with the differences in scale between different variables for each student, the cluster elements should be properly normalized. This process, which uses Eq. (1), is performed prior to the cluster analysis so as to make the original data dimensionless.

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

Where, x_{nm} , $\bar{x}_m$ and s_m represent original, average, and the standard deviation of each variable/measurement, respectively, for any particular observation.

In cluster analysis, cluster elements are grouped according to their similarities, or more specifically, the distances between them. Therefore, the smaller the distances between the elements, the more similar they are and the more likely they belong to the same cluster. For our study, squared Euclidean distance, as shown in Eq. (2), is implemented for calculating the distance between clusters.

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

Where, d_{ij}^2 is the squared Euclidean distance between state elements i and j; x_{im} is the m^{th} element in individual i; and x_{jm} is the m^{th} element in individual j.

The K-means algorithm is one of the most popular techniques in non-hierarchical clustering, which requires specifying the number of clusters arbitrarily [31]. Compared to hierarchical clustering, K-means clustering is a faster and more reliable method, especially for applications with large, high dimensional data sets. Furthermore, the K-means algorithm repeatedly reassigns elements to clusters so the same element can move from cluster to cluster during the analysis. Considering the advantages of non-hierarchical clustering, the proposed methodology for this study is to select the number of clusters desired (e.g., the teacher can easily set the cluster number to a desired number for their own scale). Then the K-means procedure is used to actually form the clusters.

4. RESULTS

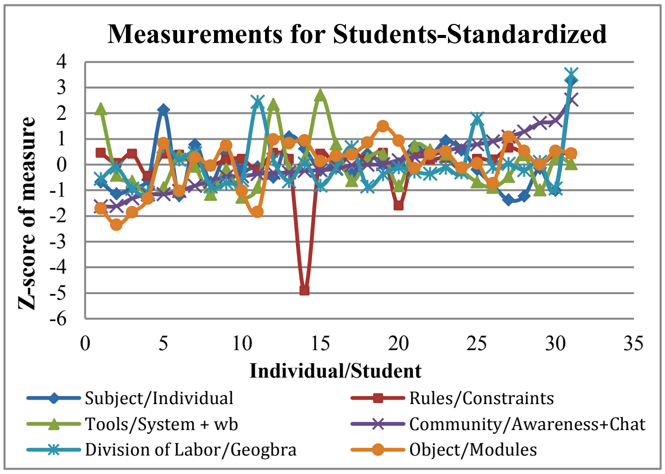

Figure 4. Standardized measurements for Students

For the purposes of analysis, this study used data from 31 users in the VMT environment. Based on the metrics construction methodology discussed above, we obtained the metric distribution shown in Figure 4. This graph shows that each metric is in fluctuation and that there is no unified trend for the six measures. For example, for the fifth and sixth students, the sixth one is higher than the fifth one in the Community value. However, in all of the other five dimensions such as Subject, Tools, Rules, etc., the fifth student surpasses the sixth one. Therefore, it is hard to assess participation in CSCL activities based on only a single metric. We need to consider the activity of learning from six dimensions (as a whole system) to assess students' performance. Thus, we employed the K-means clustering method to classify students according to their behaviors in the system across measures. This approach establishes sets of students with behavioral similarities previously invisible to teachers and students.

4.1 Clustering Results

Using the K-means algorithm, Table 2 shows the final clustering results. The values under each cluster are the mean scores for the corresponding measurements.

Table 2. K-means Clustering Results – 3 Clusters

Cluster Size		Cluster1	Cluster2	Cluster3
		11/31	12/31	8/31
Measure	Range (Standardized)	Cluster Means		
Subject	[-1.38, 3.27]	0.772	-0.238	-0.704
Tools	[-4.91, 0.65]	0.038	-0.082	0.071
Community	[-1.28, 2.71]	-0.385	0.862	-0.764
Rules	[-1.63, 2.53]	0.322	0.181	-0.714
Division of Labor	[-1.09, 3.52]	0.483	-0.354	-0.132
Object	[-2.35, 1.48]	0.438	0.450	-1.277

4.1.1 Interpretation

Interpretation of results from the cluster analysis is a matter of comparing the means of each cluster across the six measures. For instance, cluster 1 has the highest mean in the Subject, Rules and Division of Labor measures, but the lowest mean in the Community and Object measures. By referencing how each of these measures was defined in Table 1, we can interpret the participation results of students in this cluster. In future research we will explore visualization of results to make interpretation simpler for teachers. Below, we examine each measure's distribution with interpretations and discuss the characteristics of student participation in VMT activities followed by a recommendation for teachers. These clusters are ordered by cluster size.

4.1.2 Cluster 1: Personally Participative, Active, Limited Communication

This group of students has 11 members who tend to be personally participative since they have the highest value (0.772) on the Subject dimension, which is reflected by a student' personal endeavors or efforts on individual modules or tasks. Therefore, they are actively exploring the VMT environment on their own even though there is no peer pressure on them to participate in group projects. In addition, those students are also rated highest (0.483) on Division of Labor and Rules (0.322), which means each student in this cluster contributes significantly in the group projects or tasks such as geometry artifact construction using varying functions residing in the VMT environment.

The other three dimension Tools (0.038), Community (-0.385), and Object (0.438) are ranked medium across the three clusters. Considering that their measure on Division of Labor is larger than the other groups, it is easy to infer that these students are very active in some of the modules but are perhaps inactive in others. In addition, the teacher should consider encouraging those students to have more interaction with the students in their group in that they were rated relatively low in the Community measure. Still, in general, students in this cluster have the highest performance in CSCL activities.

4.1.3 Cluster 2: Collaboratively Participative, Not Deep into Learning

This cluster has 12 students in it who are not as personally participative compared to students in Cluster 1. However, they have the highest score on the Object dimension (0.450), which mainly measures how active those students were in various modules. In addition, students in this cluster have the highest value on the Community (0.862) measure indicating that those students actively communicate with their team members. Nonetheless, those students have the lowest measure on Division of Labor. This shows that even though they actively participate in conversations within their own groups, they made limited contributions to geometry object construction. This may also be reflected from their score on the Tools as the lowest -0.082 because those students mainly engaged in the chat function but do not actively employ tools and functions that the VMT environment offers.

The teacher might infer that these students are not that familiar with the VMT software environment since they are not devoted to building geometry objects using various functions (Tools). They might just master the basic functions such as using chat to communicate with other students. Therefore, the teacher could spend more time introducing the environment or preparing better introductory materials for the students in this cluster. Generally, students have a medium performance in this cluster.

4.1.4 Cluster 3: Less Participative, Great Group Learner

This cluster has 8 students who performed poorly in personal learning tasks, because for the Subject dimension, which requires personal endeavors in learning modules, they have the lowest score (-0.704). Similarly, they did not participate in the group module well: they have the lowest score on Object (-1.277), which consists mainly of group tasks and modules. Also, students in this cluster have the lowest value on the Community dimension (-0.764) reflecting that they did not engaged in communication with their team members. However, these students have the highest score on the Tools measure (0.071), which shows that they actively explore the VMT environment and used various functions. Nevertheless, these students might get too focus on the geometry construction related functions but overlook others, therefore, they are rated the lowest in the Rules (-0.714). In sum, these students are the poorest performers among the three groups.

5. DISCUSSION & CONCLUSION

Current assessments in CSCL are time consuming, or lacking in reliability, validity or individual accountability. Our work here does not claim to address those shortcomings. Instead, it is hoped that this assessment method may serve as a supplemental tool for helping to identify individual student trajectories and categories in synchronous, online learning environments like VMTwG. Specifically, this method is objective in nature because it deals with trace data of students' actions automatically recorded by the VMT environment, and presents those interactions through the lens of Activity theory. When the methodology is set, it would get the same results every time (given the same actions by the same students). In the traditional methods, different teachers may grade or assess students' performance differently, and peer assessments may be affected by various biases [18]. This methodology improves the reliability of assessment in CSCL.

The objectivity we argue exists in this methodology is not objective assessment of learning specifically. We argue that through the lens of activity theory, the research presented here demonstrates that student behavior can be compared more reliably by teachers and other students than would be the case a) without the specific approach described, b) using a less theoretically informed method or c) without an intimate knowledge of the learning activities being examined. Learning analytics, in this study, is the intersection of behavioral trace data, theory and systematic application of established methods of computing multidimensional similarity. Our aim is to support teachers and students by systematically making behavioral signals visible.

In addition, the proposed method (from measure construction to clustering results) can be totally automatic. Hence, rather than coding various interviews, observation notes, and discussions, it can be realized by a computer program which is much more efficient. Also, in the traditional assessment methods, there is often a lack of established standards and protocols in processing the qualitative interviews or chat log data [11], in turn affecting the reliability of the assessment method.

In terms of validity, learning and performance is complex in nature and it is hard to assess by single measures such as tests, quizzes or essays [2, 32]. These traditional methods are not able to incorporate learning that happens in the process and the context. Our methodology focuses on process, learning environment and context for learning and performance assessment (participation in collaborative activities). While learning continues to be evaluated principally at the individual level, systems like VMTwG focus on the small group unit of analysis. Through our systematic approach, and the continued development of approaches like ours, we think time consuming assessment and research methods can be accelerated.

Some courses or training programs (like that analyzed here) are geared more toward exploration. Foci of these courses are more prone to impart concepts or explore into a subject (like geometry) [33]. Students do not generate objects or artifacts for teachers to assess. The learning is integrated in the process. In such cases, traditional methods such as tests or evaluation of the final product might not be a good fit to measure this type of learning or performance. By contrast, our proposed method is a powerful way to assess performance and participation in the learning process. This gives teachers an insight into the holistic participation of students in collaborative online activities. Cluster results provide a window into what is going on in the activity system of individual students so that teachers can more fully address student participation in collaborative learning and use other means (such as qualitative analysis) to investigate student collaboration quality.

One of the major means of assessment in CSCL is to evaluate group tasks and or individual tasks, then assign the same grade to the group or differentiate the grade by looking at participation in individual tasks within the group. In the current educational context, which highlights individual success and effort, group grading may not be an ideal solution [2]. Our method emphasizes individual contribution and endeavors in the CSCL activities in a more objective and reliable way. If, as Stahl argues [34], group cognition is the product of interactions between group members, than the scalability of technologically mediated small group learning will require new ways of assessment at scale.

6. LIMITATIONS

The proposed method is purely quantitative. It does not consider the quality of ultimate artifacts or objects that might be generated at the end of a course or training program. In addition, communication and language is also a powerful way of learning [35, 36]. A lack of systematic assessment of the qualitative aspects of collaborative work is one of the limitations of our proposed method. For instance, researchers less familiar with the

VMTwG environment and without experience analyzing interactions in the environment would have a difficult time replicating our results in another context. Still, this method can serve as a supplemental tool to dampen or provide triangulation for evaluations of performance and outcomes from qualitative or other assessment methods of learning and performance in CSCL.

7. FUTURE WORK

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. We are going to explore natural language processing of the chat log data and incorporate it into our activity theory measure construction system in order to further inform learning assessment in CSCL. Additionally, our clustering approach puts individuals with similar behavior patterns into groups, and it would be helpful to further explore approaches that are able to identify individual behavioral distinctions. Last, because this work solely focused on a purely CSCL setting, the applicability of this methodology to other learning environments needs to be tested.

Our future studies will explore dynamic constructions of student relations and understanding of mathematical constructs through tool use; articulated previously as instrumental genesis [37]. Nardi & Kaptelinin [38] described how instrumental genesis could inform the development of activity theory as a lens for practical HCI design. In this paper, we have applied knowledge from working within a system over a course of years to a specific way of synthesizing insight from a combination of activity theory and computation; theoretically informed data mining. The double loop learning [39] and research we are aiming for will reconceptualize the design of systems, the logs those systems generate and the analysis provided back to users. Perhaps beginning with consideration of feedback during the design cycle – and a corresponding primacy of what log data to generate – will inspire a new era in learning analytics and learning technologies design.

8. REFERENCES

[1] *1st International Conference on Learning Analytics and Knowledge*, Banff, Alberta, February 27- March 1, 2011. Available from https://tekri.athabascau.ca/analytics/.

[2] Dennen, Vanessa Paz. Looking for evidence of learning: Assessment and analysis methods for online discourse. *Computers in Human Behavior* 24, no. 2 (2008): 205-219.

[3] Strijbos, J-W. "Assessment of (computer-supported) collaborative learning." *Learning Technologies, IEEE Transactions on* 4, no. 1 (2011): 59-73.

[4] Stahl, Gerry, Timothy Koschmann, and Dan Suthers. Computer-supported collaborative learning: An historical perspective. *Cambridge handbook of the learning sciences* 2006 (2006).

[5] Strijbos, Jan-Willem, Paul A. Kirschner, and Rob L. Martens, eds. *What we know about CSCL: And implementing it in higher education*. Vol. 3. Springer, 2004, 31-50.

[6] Strijbos, Jan-Willem, and Frank Fischer. Methodological challenges for collaborative learning research. *Learning and Instruction* 17, no. 4 (2007): 389-393.

[7] Prins, Frans J., Dominique MA Sluijsmans, Paul A. Kirschner, and Jan-Willem Strijbos. Formative peer assessment in a CSCL environment: A case study. *Assessment & Evaluation in Higher Education* 30, no. 4 (2005): 417-444.

[8] Wikipedia contributors, Assessment. Wikipedia, The Free Encyclopedia, http://en.wikipedia.org/wiki/Assessment (accessed Oct 1, 2013). Oct 1, 2013, 02:02 (UTC)

[9] Slavin, Robert E. Cooperative learning in teams: state of the art. *Educational Psychologist* 15, no. 2 (1980): 93-111.

[10] De Wever, Bram, Hilde Van Keer, Tammy Schellens, and Martin Valcke. Roles as a structuring tool in online discussion groups: The differential impact of different roles on social knowledge construction. *Computers in Human Behavior* 26, no. 4 (2010): 516-523.

[11] Gress, Carmen LZ, Meghann Fior, Allyson F. Hadwin, and Philip H. Winne. Measurement and assessment in computer-supported collaborative learning. *Computers in Human Behavior* 26, no. 5 (2010): 806-814.

[12] Hathorn, Lesley G., and Albert L. Ingram. Cooperation and collaboration using computer-mediated communication. *Journal of Educational Computing Research* 26, no. 3 (2002): 325-347.

[13] Kerr, Norbert L. Motivation losses in small groups: A social dilemma analysis. *Journal of Personality and Social Psychology* 45, no. 4 (1983): 819.

[14] Salomon, Gavriel, and Tamar Globerson. When teams do not function the way they ought to. *International journal of Educational research* 13, no. 1 (1989): 89-99.

[15] Ross, John A., and Carol Rolheiser. Student assessment practices in cooperative learning. *Cooperative Learning: The social and intellectual outcomes of learning in groups* (2003): 54-68.

[16] Hanrahan, Stephanie J., and Geoff Isaacs. Assessing Self-and Peer-assessment: the students' views. *Higher education research and development* 20, no. 1 (2001): 53-70.

[17] Kwan, Kam-Por, and Roberta Wong Leung. Tutor versus peer group assessment of student performance in a simulation training exercise. *Assessment & Evaluation in Higher Education* 21, no. 3 (1996): 205-214.

[18] Pond, Keith, Rehan Ul-Haq, and Winnie Wade. Peer review: a precursor to peer assessment. *Programmed Learning* 32, no. 4 (1995): 314-323.

[19] Mirriahi, Negin, and Dawson, Shane. The pairing of lecture recording data with assessment scores: A method of discovering pedagogical impact. In *Proceedings of the International Conference on Learning Analytics and Knowledge, 2013* (Leuven, Belgium April 8-12, 2013), 180-184.

[20] Koulocheri, Eleni, and Xenos, Michalis. Considering formal assessment in learning analytics within a PLE: The HOU2LEARN case. In *Proceedings of the International Conference on Learning Analytics and Knowledge, 2013* (Leuven, Belgium April 8-12, 2013), 28-32.

[21] Barab, Sasha A., Michael Barnett, Lisa Yamagata-Lynch, Kurt Squire, and Thomas Keating. Using activity theory to understand the systemic tensions characterizing a

technology-rich introductory astronomy course. *Mind, Culture, and Activity* 9, no. 2 (2002): 76-107.

[22] Engeström, Yrjö. Learning by expanding. An activity-theoretical approach to developmental research. (1987).

[23] Leont'ev, Aleksei N. The problem of activity in psychology. *Journal of Russian and East European Psychology* 13, no. 2 (1974): 4-33.

[24] Nardi, Bonnie A., ed. *Context and consciousness: Activity theory and human computer interaction*. The MIT Press, 1996.

[25] Basharina, Olga K. An activity theory perspective on student-reported contradictions in international telecollaboration. *Language Learning & Technology* 11, no. 2 (2007): 82-103.

[26] Engeström, Yrjö. Activity theory and individual and social transformation. *Perspectives on activity theory* (1999): 19-38.

[27] Howison, James, Wiggins, Andrea, and Crowston, Kevin.Validity Issues in the Use of Social Network Analysis with Digital Trace Data. *Journal of the Association of Information Systems*. 12(2), 2, (2012).

[28] Goggins, Sean P., Christopher Mascaro, and Giuseppe Valetto. "Group informatics: A methodological approach and ontology for sociotechnical group research." *Journal of the American Society for Information Science and Technology* (2013).

[29] Duda, Richard O., Peter E. Hart, and David G. Stork. *Pattern classification*. John Wiley & Sons, 2012.

[30] Guo, Rui, and Zhang Yu. Identifying Time-of-Day Breakpoints Based on Non-intrusive Data Collection Platforms. *Journal of Intelligent Transportation Systems* (2013).

[31] Xu, Chengcheng, Pan Liu, Wei Wang, and Zhibin Li. Evaluation of the impacts of traffic states on crash risks on freeways. *Accident Analysis & Prevention* 47 (2012): 162-171.

[32] Donahoe, John W., and David C. Palmer. *Learning and complex behavior*. Allyn & Bacon, 1994.

[33] Gardner, Elizabeth B., John J. Boitano, Nicholas S. Mancino, Denis P. D'Amico, and Eliot L. Gardner. Environmental enrichment and deprivation: effects on learning, memory and exploration. *Physiology & Behavior* 14, no. 3 (1975): 321-327.

[34] Stahl, Gerry, Timothy Koschmann, and Dan Suthers. "Computer-supported collaborative learning: An historical perspective." *Cambridge handbook of the learning sciences* 2006 (2006).

[35] Wells, Gordon. *Learning Through Interaction: Volume 1: The Study of Language Development*. Vol. 1. Cambridge University Press, 1981.

[36] Fromkin, Victoria. *An introduction to language*. Cengage Learning, 2013.

[37] Guin D, Trouche L. Mastering by the teacher of the instrumental genesis in CAS environments: necessity of intrumental orchestrations[J]. *Zentralblatt für Didaktik der Mathematik*, 2002, 34(5): 204-211.

[38] Kaptelinin V, Nardi B A. Acting with technology[M]. MIT Press, 2006.

[39] Argyris C, Schon D A. Theory in practice: Increasing professional effectiveness[M]. Jossey-Bass, 1974.