Identifying Profiles of Collaborative Problem Solvers in an Online Electronics Environment

Jessica Andrews-Todd
Educational Testing Service
660 Rosedale Rd.
Princeton, NJ 08540
+1(609) 734-5809
jandrewstodd@ets.org

Carol Forsyth
Educational Testing Service
90 New Montgomery, Ste. 1500
San Francisco, CA 94105
+1(415) 645-8465
cforsyth@ets.org

Jonathan Steinberg
Educational Testing Service
660 Rosedale Rd.
Princeton, NJ 08540
+1(609) 734-5324
jsteinberg@ets.org

André Rupp
Educational Testing Service
660 Rosedale Rd.
Princeton, NJ 08540
+1(609) 252-8545
arupp@ets.org

ABSTRACT

In this paper, we describe a theoretically-grounded data mining approach to identify types of collaborative problem solvers based on students' interactions with an online simulation-based task about electronics concepts. In our approach, we developed an ontology to identify the theoretically-grounded features of collaborative problem solving (CPS). After interaction with the task, students' log files were tagged for the presence of 11 CPS skills from the ontology. The frequencies of the skills were clustered to identify four unique profiles of collaborative problem solvers – Chatty Doers, Social Loafers, Group Organizers, and Active Collaborators. Relationships among cluster membership, task performance, and external ratings of collaboration provide initial validity evidence that these are meaningful profiles of collaborative problem solvers.

Keywords

Collaborative Problem Solving, Ontology, Assessment, Simulation-based Assessment, Discourse

1. INTRODUCTION

In our modern society, the nature of workplace performance has changed fundamentally through technology. An increasing number of complex tasks are being carried out in groups, often supported through digital tools with features that support collaboration. Accordingly, there has been increased attention in the assessment community on relevant competencies such as collaborative problem solving (CPS), a skill with multiple components that have been identified as important for success in the 21st century workforce [3].

Competency in CPS has been defined as "the capacity of an individual to effectively engage in a process whereby two or more agents attempt to solve a problem by sharing the understanding and effort required to come to a solution and pooling their knowledge, skills, and efforts to reach that solution" [17]. The complexity of this construct in having a cognitive dimension associated with problem solving processes and an interpersonal dimension associated with collaboration processes has made assessing CPS difficult, if not impossible, to carry out with traditional types of assessment such as multiple-choice questions with almost any sense of fidelity and generalizability [5]. As a result, there has been a turn to online learning environments such as games and simulations, which allow individuals to interact around complex problems and capture all actions and discourse in the environment as evidence of competency for assessment purposes.

While online environments offer promise for CPS assessment, there are challenges that exist. First, as with more traditional forms of assessment, assessment developers must conceptualize what skills define the construct and what actions and discourse would be indicative of those skills in the environment. Second, one must develop methods to make sense of the large streams of fine-grained data generated during real-time interaction in the environment [10].

In the current paper, we use a theoretically-grounded data mining approach [6] to discover profiles of various types of collaborative problem solvers that are strongly rooted in theory associated with collaboration, cognitive and social psychological research. Specifically, we describe the principled approach we used to conceptualize what skills make up the CPS construct, how we extracted evidence of those skills from the large streams of log data, and how we aggregated that information to create profiles that describe different types of collaborative problem solvers.

2. METHODS

2.1 Participants

Students in electronics and engineering programs were recruited from universities and community colleges across the United States. There were 129 individuals who completed the study in groups of three (i.e., 43 groups) that were randomly assembled. Of those students who reported their gender, 81% were males and 17% were females with 2% unreported. Of those who reported their race, 51% were White, 7% were Black or African American, 6% were Asian, 2% were American Indian or Alaska Native, 10% reported being more than one race, 2% reported Other, with 2% unreported. For ethnicity, 22% reported being Hispanic. The average age among students was 24 in a range of 16 to 60.

2.2 Task and Measures

Students completed a pre-survey that asked for their background information (e.g., age, gender, level of education) as well as their preferences for working in groups relative to independently and beliefs about the importance of collaboration. Instructors were then asked to randomly assemble their students into groups to complete an online simulation-based task on electronics concepts. The students worked in a computer lab and collaborated completely online in a computer-mediated environment described next.

In the task, called the Three-Resistor Activity, students worked in groups of three, each on a separate computer, and each running a fully functional simulation of a portion of an electronic circuit. The individual simulations were linked together to form a complete series circuit. The environment included a digital multimeter (DMM), two probes (red and black) from the DMM, a resistor, a calculator, a zoom button, a chat window, and a submit button (see Figure 1 for a screenshot of the task interface). These components allowed students to take measurements, view their circuit's resistance, perform calculations, zoom out to view (but not interact with) other teammates' circuits, communicate with teammates, and submit their work.

The individuals in each team were given the same task goal, which consisted of setting their resistors so that the voltage across these matched specified goal values. Since the circuits were connected in series, a change made to any one of these affected the current through the circuits and therefore the voltage drop across each of the circuits. Thus, rather than attempting to achieve the goal independently, team members needed to share information and coordinate their efforts to reach the goal voltage values across all the circuits. There were four levels of the task that increased in difficulty. At higher difficulty levels of the task, in addition to achieving their goal voltage values, the students were also asked to collaborate to determine the unknown resistance and supply voltage of an external, fourth circuit in the series. Students were allowed to communicate only using a chat window and could "zoom out" to see one another's circuits, but could only alter or make measurements on their own circuits. As students worked to achieve the goal voltages across four task levels, all of their relevant actions (e.g., DMM measurements, resistor changes, calculator entries, chat submissions) were time-stamped and logged to a database.

Table 1 provides an overview of the characteristics of each task level. Across the four task levels, the difficulty of the task increased either by presenting a more complicated problem (e.g., providing different goal voltages for each teammate in Level 2) or reducing the amount of information given (e.g., the external voltage in Levels 3 and 4). These changes increased the need for collaboration, as students were required to share more information and communicate more to identify unknown variables. Specifically, in Level 1, students were given the unknown resistance and supply voltage of an external, fourth circuit in the series and the goal voltages that needed to be reached were the same for each teammate. Having the same goal voltages for each circuit limited the amount of information that needed to be shared for each teammate to reach their goal. In Level 2, students were again given the unknown resistance and supply voltage of an external, fourth circuit in the series, but each teammate was now given a different goal voltage that they were required to reach. In Level 3, students were given the value of the resistance of the external circuit and again had different goal voltages to reach; however, the supply voltage of the external circuit was not provided. Thus, the team needed to reach the goal voltage for each circuit, but also discover and submit the supply voltage value and unit for the external circuit. In Level 4, students needed to discover and report the values and units for both the unknown resistance and the supply voltage of the external, fourth circuit as well as reach the specified and different goal voltages on each teammate's circuit.

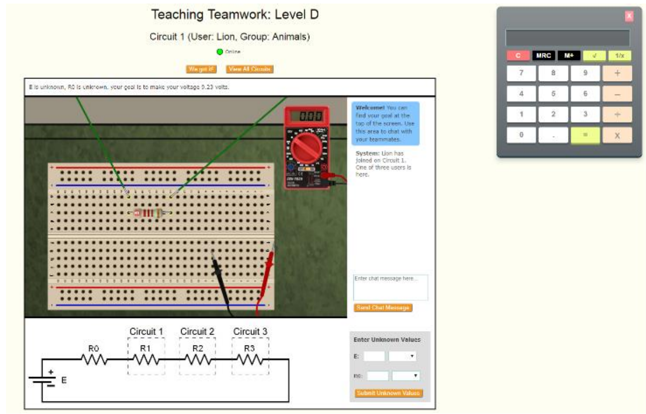

Figure 1. Screenshot of the Three-Resistor Activity.

Table 1. Overview of Task Levels

Task Level	External Voltage (E)	External Resistance (R0)	Goal Voltages
1	Known by all teammates	Known by all teammates	Same for all teammates
2	Known by all teammates	Known by all teammates	Different for each teammate
3	Unknown by teammates	Known by all teammates	Different for each teammate
4	Unknown by teammates	Unknown by teammates	Different for each teammate

2.3 Competency Model

A CPS ontology (similar to a concept map) was developed to conceptualize the CPS construct. It provides a theory-driven representation of the targeted skills and their relationships, linking the skills to observable behaviors in the electronics task that would provide evidence of each skill. The top level of the ontology provides generalizable construct definitions for CPS (e.g., sharing information as one skill associated with the construct) that can be implemented in other work seeking to assess CPS or other related constructs. This top layer was developed based on an extensive literature review of CPS frameworks and other related research areas such as computer-supported collaborative learning, organizational psychology, individual problem solving, and linguistics [9, 12, 14, 15, 16, 17, 18, 22]. Each lower layer of the ontology becomes more specific describing CPS as interpreted within a domain (e.g., sharing status updates) and then within the task environment in the domain (e.g., sharing the status of the resistance in a circuit). Links between the layers describe how behaviors at lower levels can be combined to make inferences about cognitive behaviors at higher levels. In our research, the ontology designated the lower level features corresponding to over-arching social and cognitive dimensions. These lower level features were then extracted from log files prior to analysis. Figure 2 shows the structure for a portion of the CPS ontology with nodes corresponding to high-level CPS skills, sub-skills, features, and observable variables that can be inferred from the features, along with links indicating the relationships between the nodes.

Figure 2. CPS ontology fragment structure.

The full ontology has nine high-level skills associated with CPS that we sought to identify in the data. Four skills correspond to the social dimension of CPS (i.e., maintaining communication, sharing information, establishing shared understanding, negotiating) and five skills correspond to the cognitive dimension of CPS (i.e., exploring and understanding, representing and formulating, planning, executing, monitoring). Maintaining communication corresponds to content irrelevant social communications [12]. This includes general off-topic communication (e.g., discussing what was eaten for breakfast), rapport building communication (e.g., greeting or praising teammates), and inappropriate communication (e.g., cursing). Sharing information corresponds to content relevant information communicated during collaboration. This includes the sharing of one's own information (e.g., sharing information related to the status of one's own work during the task), sharing task or resource information (e.g., communicating what tools are available in the task environment), and sharing understanding (e.g., sharing metacognitive information about the state of one's understanding). Establishing shared understanding corresponds to communicators attempting to learn the perspectives of others as well as trying to establish that what has been said is understood [4, 17]. This skill would include requesting information from teammates to verify that everyone has a common understanding, providing responses to teammates that verify comprehension of another's contribution, and making repairs when problems in shared understanding arise. Negotiating refers to communication that identifies whether or not conflicts exist in the ideas among teammates and seeks to resolve those conflicts when they arise [9]. This skill includes expressing both agreement and disagreement, and attempting to reach a compromise.

For the cognitive dimension, exploring and understanding refers to actions taken to build a mental representation of pieces of information associated with the problem. This includes interacting with the task environment to explore the problem space and demonstrating understanding of given information and information acquired while interacting with the environment. Representing and formulating refers to actions and communication in the service of building a coherent mental representation of the whole problem space. This includes developing a verbal or graphical representation of the problem and formulating hypotheses [17]. Planning corresponds to communication around developing a plan or strategy to solve the problem. This includes determining the overall goal, setting sub-goals or steps to carry out, and developing and revising strategies [9, 17]. Executing corresponds to actions and communication used in the service of carrying out a plan. This includes taking actions to enact a strategy, making suggestions for actions a teammate should carry out, and communicating to teammates the actions one is taking to carry out the plan. Monitoring refers to actions and communication associated with monitoring progress toward the goal and monitoring the organization of the team [16, 17]. This includes communicating one's own progress toward the goal, checking on the progress of teammates, and determining whether teammates are present and following the rules of engagement or their roles in completing tasks.

2.4 Qualitative Coding

The CPS ontology was used to create a rubric for raters to carry out qualitative coding of the log data to identify evidence of high-level CPS skills from low-level student discourse and actions. The nodes and links corresponding to each CPS skill in the ontology were transformed into extensive written protocols that included the high-level CPS skills, any sub-skills associated with the high-level skills, definitions for skills and sub-skills, example behaviors from the log data that would be indicative of each skill, and the action types associated with each skill (e.g., chat, calculation, measurement, submit). Two raters coded the content of students' discourse and their actions for the display of nine CPS skills. Evidence for two of the nine high-level CPS skills from the ontology could be found in both chats and actions (i.e., monitoring and executing) and were thus split into separate action and chat skills. As a result, the 11 coded skills were maintaining communication, sharing information, establishing shared understanding, negotiating, exploring and understanding, representing and formulating, planning, executing actions, executing chats, monitoring actions, and monitoring chats. Coding was done at the level of each log file event (i.e., each action submission or submission of a chat [utterance level] even if sequences of utterances mapped onto a singular CPS skill). Each of the 20,947 log file events only received one code. The inter-rater reliability between the two raters was high (Kappa = .84) based on a randomly selected sample of 20 percent of the data (approximately 4,200 events) that were double-coded.

On the social dimension, for maintaining communication, raters examined the log data for evidence of off-topic communication (e.g., "I should have drank coffee this morning"), rapport building communication (e.g., using chat emoticons, greeting teammates, apologizing, praising teammates), and inappropriate communication such as curse words or messages that degrade teammates (e.g., "you're an idiot"). For sharing information, raters looked for evidence of individuals sharing their own information for the problem (e.g., sharing what circuit board they were on, their goal voltage values, or resistance values on their board), sharing task or resource information (e.g., sharing where the zoom button was located, sharing that there was a calculator to use in the environment), and sharing their understanding (e.g., metacognitive statements such as "I don't get it"). For establishing shared understanding, raters looked for evidence of individuals requesting information from their partners (e.g., "what is your resistance?"

"what values do we need?"), and providing responses that indicate comprehension or lack of comprehension of a teammate's statement (e.g., "ok," "I hear you," or requests for clarification). For negotiating, raters looked for evidence of individuals expressing agreement (e.g., "You are right"), expressing disagreement (e.g., "that's not right"), and revising their own ideas or proposing alternate ideas.

On the cognitive side, raters looked for evidence of exploring and understanding by identifying actions in which individuals unsystematically made changes to task components in an effort to explore the interface. Unsystematic actions were defined as seemingly exploratory actions that were taken prior to developing a plan (e.g., spinning the dial on the digital multimeter, changing the resistance values several times in a few seconds). For representing and formulating, raters looked for evidence of individuals verbally communicating what the problem was (e.g., "this is a series circuit") and communicating hypotheses for how their actions would affect the environment. For planning, raters looked for evidence of individuals communicating goals (e.g., "We need 6.69 volts across our resistors") and communicating strategies to their teammates (e.g., "ok we set our values to R and find current"). For executing actions, raters looked for actions that individuals took to carry out the plan or strategy (e.g., changing their voltage values to the voltage suggested by a teammate or performing a calculation associated with Ohm's Law). For executing chats, raters looked for evidence of individuals making suggestions or directing their teammates to perform actions associated with their plan (e.g., "Adjust yours to 300 ohms") and reporting their own actions that they were taking to carry out the plan (e.g., "Let me go a little lower and then readjust"). For monitoring actions, raters looked for evidence of individuals carrying out actions associated with monitoring the team's progress toward the goal (e.g., clicking the submit button to receive feedback about success in solving the problem) or monitoring teammates (e.g., using the zoom feature to view the state of a teammate's circuit board). For monitoring chats, raters looked for evidence of individuals stating the result of their monitoring of progress toward the goal (e.g., "I've got my goal voltage"), monitoring the status of teammates (e.g., "Where is Rain?"), and prompting teammates to perform tasks (e.g., "Let's get a move on Sleet").

3. ANALYSES AND RESULTS

The analyses were conducted in two stages. First, the frequencies of the 11 CPS skills displayed by each individual were clustered with a hierarchical approach to discover meaningful profiles. Second, the profiles were validated by their relationship to performance and self-report measures with non-parametric inferential statistical tests and Monte Carlo simulations due to the abnormal distributions of the variables.

3.1 Cluster Analysis and Profiles

We chose an exploratory clustering method [21] for uncovering potential profiles of collaborative problem solvers in part because we had no formal a priori theory regarding the number and composition of these profiles. Additionally, as the sample size (N=129) did not warrant methods like K-means which are typically applied to larger samples [13], Ward's Method was employed to cluster the frequencies of each CPS skill displayed to allow us to examine the breakdown of possible clusters so that a meaningful number of clusters could be chosen. The final number of clusters was determined based on an initial interpretation of the theory stated in existing literature in collaboration and psychological research. Thus, these are preliminary findings and to date no gold standard exists for the collaborative problem solving domain.

A four-cluster solution was most defensible from a theoretical perspective and the expected relationships to other variables that resulted which will be explained in later sections; Table 2 shows the frequencies for this solution. Specifically, the learners in the four clusters differed systematically in the frequencies of CPS skills that were displayed. The four clusters were named Chatty Doers, Social Loafers, Group Organizers, and Active Collaborators. In the next section, we describe the key behavioral patterns in each cluster based on CPS skill frequencies standardized to the total sample and discuss the relevant theory explaining the type of collaborative problem solver that may display the patterns of behavior.

Table 2. Collaborative Problem Solver Profiles

Profile	Frequency	Percent of Sample
Chatty Doers	35	27.1
Social Loafers	68	52.7
Group Organizers	16	12.4
Active Collaborators	10	7.8

3.1.1 Chatty Doers

Students in Cluster 1, labeled "Chatty Doers" (n=35) were high ($z \geq 0.20$) on executing actions and maintaining communication, somewhat high ($0.10 \leq z < 0.20$) on planning and sharing information, and were low ($z \leq -0.20$) on monitoring actions. These students were labeled "Chatty Doers" due to their high levels of maintaining communication chats and executing actions. Chats associated with maintaining communication were communications that were social in nature, but not relevant to solving the problem [12]. These included discussing what one did last week, discussing homework from the night before, and praising teammates. Thus, these individuals were designated as chatty more generally given their off-topic, social communication that was absent of high levels of communication related to skills such as negotiating or establishing shared understanding. These individuals also engaged in a high level of executing actions relative to other individuals which included making resistor changes and performing calculations. Thus, these individuals were the doers carrying out many of the actions associated with executing the team's plan.

3.1.2 Social Loafers

The standardized means for Cluster 2, labeled "Social Loafers" (n=68) displayed below average demonstration ($z < 0.00$) of almost all skills. These students were named "Social Loafers" given their low levels of the CPS skills which may be explained by a social psychological phenomenon in which individuals decrease their individual effort when working in groups [11] as they each assume another member will take the lead in solving the problem. Students in this cluster appeared to do just this as they engaged in fewer collaborative problem solving behaviors relative to other individuals.

3.1.3 Group Organizers

The standardized means for Cluster 3, labeled "Group Organizers" (n=16) showed high demonstration ($z \geq 0.20$) of monitoring actions, representing and formulating, and negotiating, somewhat high demonstration ($0.10 \leq z < 0.20$) of executing chats and sharing information, and low demonstration ($z \leq -0.20$) of planning. These students were named "Group Organizers" due to their high levels of communications and actions associated with establishing and maintaining organization for the problem and the group [17]. This included things such as monitoring behaviors like using the zoom feature to monitor the state of teammates' behaviors and circuit boards, verbally representing the problem for teammates, and communicating important information to group members such as what actions are being taken to solve the problem, all of which can be in the service of keeping the group organized.

3.1.4 Active Collaborators

The students in Cluster 4, referred to as the "Active Collaborators" (n=10) showed above average demonstration ($z > 0.00$) of almost all skills, though they demonstrated low levels ($z \leq -0.20$) of maintaining communication. Cluster 4 students were named "Active Collaborators" given their high levels of almost all of the social and cognitive processes associated with CPS [8].

3.2 CPS Skill Profile Validation

The CPS skill profiles were validated by relating the cluster membership assignment to performance metrics from the task and scores from student self-reports of preference in working with others. Prior empirical studies suggest a positive relationship between demonstration of collaborative behaviors and performance outcomes [1, 8], thus we hypothesized that students demonstrating more of the skills associated with CPS would have greater success on the task as measured by the number of levels completed in the task. Number of task levels completed was treated as an individual performance measure, though contributions of other teammates could impact the score. In regard to self-report measures, we were unsure as to whether students would accurately report whether or not they thought they were good collaborators but suspected they would answer more honestly as to whether or not they preferred to work alone, thus the latter question was asked to students along with their perceptions of how important collaboration is in the real world. The cluster membership assignment, the performance metrics, and the self-ratings were submitted to Kruskal-Wallis tests with a Monte Carlo simulation to determine the significance of the relationships among the variables.

3.2.1 Cluster Membership and Performance

There was a significant relationship between cluster membership and success on the task levels (i.e., number of task levels completed) ($X^2(3,126) = 6.93$, $p < .05$ with a one-tailed test, *partial* $\eta^2 = .053$). The Monte Carlo simulation with 10,000 test samples revealed a p value of .032 (lower bound = .023; upper bound = .036). The mean ranks of the different groups based on completed task levels showed patterns in line with our prediction. Specifically, the Active Collaborators had the highest mean rank of 93.95 whereas the Social Loafers had the lowest mean rank of 61.65. Chatty Doers and Group Organizers fell in between these two groups with mean ranks of 63.89 and 63.59, respectively. Post hoc comparisons with a Bonferroni correction revealed that there was a significant difference between the Social Loafers and Active Collaborators ($p = .027$) and a marginally significant difference between the Chatty Doers and Active Collaborators ($p = .063$) in terms of mean rank of performance. All other comparisons were not significant. These results make sense as we would expect the Active Collaborators to be the high performers given that they demonstrated high frequencies of all of the necessary attributes that we had identified for effective collaborative problem solvers. It also makes sense that Social Loafers performed the poorest as these individuals demonstrated lower incidences of CPS skills.

After confirming that there was indeed a significant difference in the relationship between performance and type of collaborative problem solver, we moved on to compare cluster membership to self-reported collaboration preferences.

3.2.2 Cluster Membership and Collaboration Preferences

Recall that students completed a pre-survey that included questions about their preferences in working with others and how much they valued collaboration in the real world. We explored how responses to these questions were related to cluster membership. There was a marginally significant relationship between cluster membership and response to the question about whether or not students preferred to work alone (X^2 (3,126) = 7.23, $p = .065$ with a two-tailed test, *partial* $\eta^2 = .055$). The Monte Carlo simulation revealed a p value of .064 (lower bound = .057; upper bound = .070). The mean ranks for responses - where higher numbers indicate stronger preference to work alone - were as follows: Social Loafers (71.05), Chatty Doers (54.90), Group Organizers (54.38), and Active Collaborators (47.10). The direction of these results are consistent with what would be expected. Social Loafers who demonstrate few CPS skills and seem to expend little effort during collaborative activity would be expected to prefer to work alone. Conversely, Active Collaborators who demonstrate high incidences of CPS skills and are thus active during collaborative activity would be expected to have a preference to work with others. Chatty Doers and Group Organizers who display CPS skills, but not to the extent of Active Collaborators would be expected to fall in between the Active Collaborators and Social Loafers.

The students were also asked about their ratings as to how important collaboration is to the real world. Cluster membership had a non-significant relationship to responses on this question ($p = .465$). The mean ranks where higher numbers indicate higher importance for collaboration in the real world were as follows: Group Organizers (71.94), Chatty Doers (68.82), Active Collaborators (62.90), and Social Loafers (59.82). One possible explanation for this finding is that instructors likely informed students about the importance of collaboration in setting up the study activity so student responses may have been influenced by this information. The mean ranks were relatively high for all groups so this explanation may be appropriate, but further testing is necessary to draw any strong conclusions.

4. CONCLUSIONS

Many methods exist for discovering profiles of how students collaborate during problem solving (for a review see [7]). In the current study, we used a frequency-based cluster approach to discover cluster profiles, following a previously established approach [8]. This approach was chosen because we are discovering profiles of types of collaborative problem solvers in a discovery learning environment. That said, we acknowledge that other approaches could be considered, though they may not be the best fit in the given context. For example, for an analysis of CPS in an international assessment context [17], students interacted with a constrained environment (e.g., a dropdown menu for chat choices) making it possible for traditional psychometric approaches to sufficiently analyze the student responses and communication. Conversely, in previous research on serious games with collaboration, an Epistemic Network Analysis (ENA) approach has been used to analyze how students connect knowledge and skills during collaboration over time [19]. However, the focus of our investigation is on collaboration without including domain knowledge, though we plan on augmenting the ENA approach for our purposes in future analysis. Additional approaches focusing on group dynamics [e.g., 20] were not chosen as the goal of this investigation was to analyze student collaboration on an individual level. Therefore, we are not stating that our educational data mining approach is the only means to analyze CPS skills, but rather that it may be most appropriate for profiling individual students for CPS skills without including domain knowledge or group dynamics.

In our implementation of the frequency-based cluster approach, we demonstrated that meaningful results can emerge from incorporating theory into the approach to identify types of collaborative problem solvers. Specifically, the current approach yielded four types, namely, Chatty Doers, Social Loafers, Active Collaborators, and Group Organizers in our assessment context. The Chatty Doers displayed high levels of maintaining communication chats, or content irrelevant, social communication, and high levels of executing actions in the service of solving the problem. The Social Loafers were characterized by low levels of CPS skills in general whereas Active Collaborators were characterized by high levels of all CPS skills except maintaining communication. Group Organizers were categorized by CPS skills associated with establishing and maintaining organization for the problem and the group. Over half of the students demonstrated behaviors characteristic of Social Loafers while few students were characterized as Active Collaborators.

The profiles showed expected relationships with performance. Specifically, the Active Collaborators showed the highest levels of performance whereas the Social Loafers showed the lowest levels of performance. The performance of Chatty Doers and Group Organizers fell in between these groups. These results are consistent with prior work showing positive social and cognitive behaviors benefiting performance outcomes [8] and non-collaborative behaviors hurting performance outcomes [2]. The four cluster profiles also showed a marginally significant relationship with a self-report measure of whether or not students preferred to work with others. Social Loafers had the highest ratings of preferring to work alone perhaps because these students are less willing to expend the effort needed to sustain collaborative relationships to work with others as compared to their peers. Conversely, the Active Collaborators preferred to work with others more than did other students. This makes sense as these students are active during collaboration and thus likely willing to expend the effort needed to work with others to solve problems.

Perhaps the most important feature of this study is not necessarily the profiles themselves but rather the blending of theory with educational data mining techniques. All features of CPS were defined a priori based on a theoretically-grounded ontology with multiple levels and two dimensions of social and cognitive skills. In total, this ontology defines nearly 51 features. This method may be helpful in discovering meaningful relationships between variables in large log files from games and simulations. Furthermore, the number of clusters was defined based on theoretical grounding. We deemed the method successful based on the meaningful profiles discovered and preliminary relationships to external measures, all of which can be explained by psychological research. In the current paper, we coded high-level CPS skills based on low-level student behaviors. In future work, we intend to code at a lower, sub-skill level and incorporate methods to aggregate to higher levels in the ontology. Due to the time-intensive nature of human coding with these kind of data, we further plan to explore the possibility of automating the coding of chat data using machine learning algorithms.

There are some limitations to this study. One involves the small number of participants compared to the number of CPS skills we were attempting to measure. Additionally, we had few items to use as external correlates to our cluster profiles. In follow-up research, we are currently conducting a study with a larger sample to confirm the existence of the profiles discovered in this study and administering multiple well-constructed external measures that can potentially help build a validation argument for any discovered profiles. Another limitation of this study is that the measure used for performance outcomes incorporated the contributions of group members. As we are investigating CPS on an individual level, it would be ideal to compare student skills on an individual level to a performance measure for each individual. Thus, in an upcoming study, we have also incorporated a measure of performance that may more closely resemble individual performance but complete exclusions of group dynamics is difficult in the given environment. Thus, follow-up analyses on the group dynamics and composition are currently underway.

The current study provides preliminary results that will greatly inform the work on the upcoming data collection. Furthermore, the current study views collaboration through the lens of the Three-Resistor Activity; however, our intention is to draw upon a wide variety of tasks and content areas in upcoming studies. This future work will allow us to explore the generalizability of the CPS ontology, as its structure allows for decoupling it from content and modifying lower-level nodes to support features in other tasks.

Overall, the study demonstrates a methodology that incorporates well-detailed theory and measures emerging from the learning sciences and blends it with educational data mining. This approach resulted in meaningful profiles constructed from features defined a priori, and can serve as an example for how to combine theory and data-driven approaches to make meaningful inferences about students' knowledge, skills, and abilities from interactions in an online environment.

5. ACKNOWLEDGMENTS

This material is based upon work supported by the National Science Foundation under Grant DUE 1535224 awarded to the first author. The opinions expressed are those of the authors and do not necessarily represent views of the National Science Foundation.

6. REFERENCES

[1] Andrews, J. J. and Rapp, D. N. 2015. Benefits, costs, and challenges of collaboration for learning and memory. *Translational Issues in Psychological Science*, 1, 2, 182-191.

[2] Andrews, J. J., Kerr, D., Mislevy, R. J., von Davier, A. A., Hao, J., & Liu, L. (2017). Modeling collaborative interaction patterns in a simulation-based task. *Journal of Educational Measurement*, *54*(1), 54–69.

[3] Burrus, J., Jackson, T., Xi, N., and Steinberg, J. 2013. *Identifying the most important 21st century workforce competencies: An analysis of the occupational Information network (O* NET).* ETS Research Report RR-13-21. Educational Testing Service, Princeton, NJ.

[4] Clark, H. H. 1996. *Using language*. Cambridge University Press, New York, NY.

[5] Davey, T., Ferrara, S., Shavelson, R., Holland, P., Webb, N., and Wise, L. 2015. *Psychometric considerations for the next generation of performance assessment*. Educational Testing Service, Princeton, NJ.

[6] Forsyth, C.M., Graesser, A. C., Pavlik, P., Millis, K., and Samei, B. 2014. Discovering theoretically grounded predictors of shallow vs. deep- level learning. In Stamper, J.,

Pardos, Z., Mavrikis, M., McLaren, B.M., Eds. In *Proceedings of the 7th International Conference on Educational Data Mining (EDM 2014).* International Educational Data Mining Society, 229-232.

[7] Graesser, A.C., Cai, Z., Hu, X., Foltz, P.W., Greiff, S, Kuo, B-C. Liao, C-H. , and Shaffer, D.W. 2017. Assessment of collaborative problem solving. In R. Sottilare, A. Graesser, X. Hu, & G. Goodwin, Eds. *Design Recommendations for Intelligent Tutoring Systems: Volume 5 – Assessment*, U.S. Army Research Laboratory, Orlando, FL, 275-285.

[8] Herborn, K., Mustafić, M., and Greiff, S. 2017. Mapping an experiment-based assessment of collaborative behavior onto collaborative problem solving in PISA 2015: A cluster analysis approach for collaborator profiles. *Journal of Educational Measurement,* 54, 1, 103–122.

[9] Hesse, F., Care, E., Buder, J., Sassenberg, K., and Griffin, P. 2015. A framework for teachable collaborative problem solving skills. In P. Griffin and E. Care, Eds. *Assessment and teaching of 21st century skills.* Springer, New York, NY, 37–56.

[10] Kerr, D., Andrews, J. J., and Mislevy, R. J. 2016. The in-task assessment framework for behavioral data. In A. A. Rupp and J. P. Leighton, Eds. *The handbook of cognition and assessment: Frameworks, methodologies, and applications*. Wiley-Blackwell, Hoboken, NJ, 472-507.

[11] Latané, B., Williams, K., and Harkins, S. 1979. Many hands make light the work: The causes and consequences of social loafing. *Journal of Personality and Social Psychology,* 37, 6, 822-832.

[12] Liu, L., von Davier, A. A., Hao, J., Kyllonen, P., and Zapata-Rivera, J.-D. 2015. A tough nut to crack: Measuring collaborative problem solving. In Y. Rosen, S. Ferrara, and M. Mosharraf, Eds. *Handbook of research on computational tools for real-world skill development*, IGI-Global, Hershey, PA, 344–359.

[13] MacQueen, J. B. 1967. Some methods for classification and analysis of multivariate observations. In *Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability.* University of California Press, Berkeley, CA, 281-297.

[14] Meier, A., Spada, H., and Rummel, N. 2007. A rating scheme for assessing the quality of computer-supported collaboration processes. *International Journal of Computer-Supported Collaborative Learning,* 2, 1, 63–86.

[15] Morgan, B. B., Salas, E., and Glickman, A. S. 1993. An analysis of team evolution and maturation. *Journal of General Psychology*, 120, 3, 277–291.

[16] OECD. 2013a. *PISA 2012 assessment and analytical framework: Mathematics, reading, science, problem solving and financial literacy*. OECD Publishing, Paris.

[17] OECD. 2013b. *PISA 2015 collaborative problem solving framework*. OECD Publishing, Paris.

[18] O'Neil, H. F., Chuang, S., and Chung, G. K. W. K. 2003. Issues in the computer-based assessment of collaborative problem solving. Assessment in Education: *Principles, Policy & Practice,* 10, 3, 361–373. DOI: https://doi.org/10.1080/0969594032000148190

[19] Shaffer, D. W., Hatfield, D., Svarovsky, G., Nash, P., Nulty, A., Bagley, E. A., Frank, K., Rupp, A.A., and Mislevy, R. J. 2009. Epistemic Network Analysis: A prototype for 21st century assessment of learning. *The International Journal of Learning and Media,* 1, 1, 1–21.

[20] Von Davier, A. and Halpin, P. 2013. *Collaborative problem solving and the assessment of cognitive skills: Psychometric considerations.* ETS Research Report RR-13-41. Educational Testing Service, Princeton, NJ, 1-42.

[21] Ward, Jr., J. H. 1963. Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association,* 58, 301, 236-244. DOI: 10.2307/2282967

[22] Zhuang, X., MacCann, C., Wang, L., Liu, L., and Roberts, R. D. 2008. *Development and validity evidence supporting a teamwork and collaboration assessment for high school students.* ETS Research Report RR-08-50. Educational Testing Service, Princeton, NJ.