Are You Really A Team Player? Profiling of Collaborative Problem Solvers in an Online Environment

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

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

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

ABSTRACT

Collaborative problem solving (CPS) is considered a necessary skill for students and workers in the 21st century as the advent of technology requires more and more people to frequently work in teams. In the current study, we employed theoretically-grounded data mining techniques to identify four profiles of collaborative problem solvers interacting with an online electronics task. The profiles were created based on 11 theoretically-grounded CPS skills defined a priori. The resulting four profiles correlated in expected directions with in-task performance and had interesting relationships with external measures associated with prior knowledge and CPS skills. These results inform and partially replicate findings from our previous research using a similar approach on a smaller dataset. Implications and comparisons between the two studies will be discussed.

Keywords

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

1. INTRODUCTION

With the increasing need for technology in workplace contexts, collaborative problem solving (CPS) is considered an important 21st century skill as workers are often required to complete complex tasks in teams to solve complicated, often technical problems. Accordingly, the need to teach and assess CPS has gained increased attention by researchers [4,5]. In research seeking to assess or teach CPS skills, researchers often employ digital technologies to capture evidence and improve assessment of CPS, as this skill is complex and includes many facets.

In defining the facets of CPS there is little dispute that the construct includes social and cognitive dimensions [1,22]. The social dimension is meant to be interpersonal, including features such as sharing information and perspective taking. These types of features are associated with building a shared understanding among team members which is essential for building common ground, an important component of completing a task [6]. The cognitive dimension includes components such as planning, representing the problem, and formulating hypotheses. These components are complex in nature and therefore difficult to assess with traditional assessments such as multiple-choice questions without compromising fidelity and generalizability [7]. Therefore, assessment researchers have turned to online environments, including games and simulations, which allow for collaboration among team members to capture the discourse and complex actions necessary to evaluate CPS competency.

To evaluate CPS competency in online environments, both a competency model and advanced analytic techniques are often needed. Specifically, a competency model is necessary to identify skills and features aligning to specific constructs. Analytic approaches are needed to deal with the large streams of data stored in log files while also accounting for the underlying competency model and theoretical explanations [12].

Accordingly, in an effort to assess students' CPS skills in an online environment, we employ a theoretically-grounded data mining approach [9] incorporating a conceptual model and machine learning approaches in an iterative process. Specifically, we define a competency model based on existing literature that identifies features a priori. Our competency model is based on our prior work [1,2,3] and used to extract features in a meaningful way, and machine learning algorithms are used to profile students. We then interpret and refine algorithms based on theoretical interpretations. Thus, the process is a collaborative effort between computer scientists, learning scientists, psychometricians, and cognitive psychologists. In the current study, this principled process is used to replicate findings from a previous study [2] by discovering profiles of collaborative problem solvers that are strongly grounded in theory associated with cognitive and social psychological research. We then validate these profiles with external measures and compare these profiles with our previous findings from students interacting with the same online collaborative electronics task.

2. METHODS

2.1 Participants

Students in electronics, engineering, and physics programs were recruited from universities and community colleges across the United States to complete the study. In total, there were 378 students who participated. Of those students who reported their

Carol Forsyth, Jessica Andrews-Todd and Jonathan Steinberg "Are You Really a Team Player?: Profiles of collaborative problem solvers in an online environment" In: *Proceedings of The 13th International Conference on Educational Data Mining (EDM 2020)*, Anna N. Rafferty, Jacob Whitehill, Violetta Cavalli-Sforza, and Cristobal Romero (eds.) 2020, pp. 403 - 408

gender, 76% were males and 21% were females with 3% other, preferring not to respond, or unreported. Of those who reported their race, 62% were White, 7% were Black or African American, 8% were Asian, 10% reported being more than one race, 1% reported Other, with 4% preferring not to answer or unreported. For ethnicity, 7% reported being Hispanic. The modal age range among students was 18 to 20 years old.

2.2 Tasks and Measures

To complete the study students first completed a pretest about electronics concepts to gather information about their content knowledge, next progressed to the online electronics task, and then completed self-report measures where they rated themselves and their teammates on CPS capabilities along social and cognitive dimensions. We will first discuss the external measures and then the online electronics task.

2.2.1 External Prior Knowledge Test

The external prior knowledge test was created by a group of experts concerning the series circuit problem. First a conceptual map of the problem was created. Then, a q matrix defining skills and complexity was devised to create an equal number of questions for each electronics skill necessary to solve the series circuit problem. Next, the final items were validated by experts as well as through psychometric analysis. As a result, the original test included 28 items but only 23 saliently reflected the original intent of the test developers based on a CFA [24]. Thus, the total score per student for the 23 items is the measure of prior knowledge.

2.2.2 CPS Inventory

The CPS Inventory serves as a self-report measure of CPS skills that aligns to a competency model of CPS (which will be discussed in more detail in the next section). The Inventory consists of 14 items, seven of which correspond to social CPS behaviors (e.g., I tried to establish a good relationship with my teammates) and seven of which correspond to cognitive CPS behaviors (e.g., I helped develop a plan to solve the problem). There is a “self” version of the Inventory where participants rate their own CPS behaviors on a 4-point Likert scale (1=strongly disagree, 4=strongly agree) and a “team” version where participants rate their team’s CPS behaviors as a whole on a 4-point Likert scale (1=strongly disagree, 4=strongly agree). The CPS Inventory was administered after students completed the electronics task described next.

2.2.3 Three-Resistor Activity

Students solved a collaborative problem on electronics concepts associated with Ohm’s Law and Kirchhoff’s Voltage Law. Each student in a team of three worked on a separate computer, each running a simulation of an electronics circuit. Each student’s circuit was connected to form a series circuit.

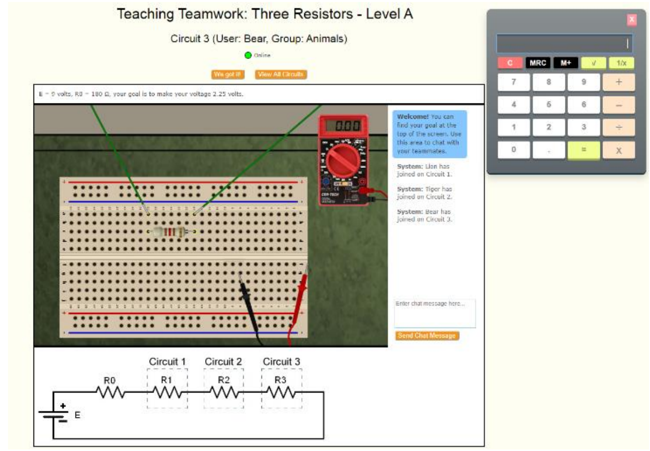

Figure 1. Screenshot of Three-Resistor Activity

Students were randomly assigned into teams by their instructors and team members were anonymized with provided usernames (e.g., Lion, Tiger, Bear). Within the interface, each student had a chat window, a digital multi-meter (DMM), probes extending from the DMM (red and black), a resistor, a zoom button, a calculator, and a submit button (See Figure 1 for a screenshot of the task interface). These features of the interface made it possible for students to communicate with teammates, take measurements, view and change the resistance on their boards, zoom out to view other teammates’ boards, perform calculations, and submit answer choices.

In the task, students had the goal of reaching a specified goal voltage value on each of their circuit boards. Because each of the circuits were connected in series, any changes made on one circuit board would affect readings on all teammates’ circuits which required the need for collaboration around coordinating actions so that everyone could reach their goal voltage values. The task has four levels which increase in difficulty as one variable changes in the task interface. In Level 1 each student had the same goal voltage value to achieve and the values of the resistance (R0) and supply voltage (E) of an external, fourth circuit in the series that students could not control were provided. In Level 2 the resistance and supply voltage of the external, fourth circuit were still provided, but each teammate now had a different goal voltage to reach. In Level 3, teammates had different goal voltages, the external resistance was provided, but the external supply voltage was unknown and needed to be found to solve the problem. In Level 4, teammates again had different goal voltages, but now both the external resistance and supply voltage were unknown and needed to be found. The task was designed so that students may only proceed to attempt the next level after completing the previous level. Therefore, levels attempted can be used as a proxy measure for performance on the electronics task. To identify the CPS skills exhibited while solving the Three-Resistor Activity, a CPS conceptual framework outlined in the form of an ontology was created, as described next.

2.2. CPS Conceptual Framework

A CPS ontology (similar to a concept map) was created using the In-Task Assessment Framework (I-TAF) approach [3,12]. This approach is an augmented version of evidence-centered design (ECD) that supports identification of features of complex constructs in online environments.

Creating the ontology required iterative refinement with the support of subject matter experts and data. The ontology was created based on literature from areas such as computer-supported collaborative learning, organizational psychology, individual problem solving, and linguistics [11,14,15, 18,19,20,21,22,23]. Data collected from the Three-Resistor Activity then informed changes to the ontology so that it most accurately reflects the construct as well as associated skills, strategies, tactics, and features based on real data collected from students interacting with the task.

To visually display the various components of CPS, the ontology is designed hierarchically. The construct (i.e., CPS) sits at the top with the two dimensions of CPS (social and cognitive) as second layer nodes. The social and cognitive nodes are linked to CPS skills associated with each dimension. Specifically, there are four skills in the social dimension and five skills in the cognitive dimension. The social dimension includes maintaining communication, sharing information, establishing shared understanding, and negotiating. The cognitive dimension includes exploring and understanding, representing and formulating, planning, executing, and monitoring. For a more in- depth discussion of this work, please refer to [1,2,3].

The nine high-level CPS skills are linked to 23 sub-skills on the fourth and fifth layers of the ontology. These sub-skills more explicitly define each of the nine CPS skills. For example, the sharing information CPS skill includes three sub-skills, sharing one's own information, sharing task or resource information, and sharing understanding. The sub-skills are connected to an evidence model which provides nodes corresponding to strategies or behaviors needed to indicate evidence of each sub-skill. The strategy nodes are then linked to tactic nodes which correspond to in-task affordances available to carry out a given strategy and subsequently feature nodes that can be inferred from individuals' behaviors. These features are identified in the log files for extraction and additional analysis. See Figure 2 for an example of the structure of a portion of the CPS ontology.

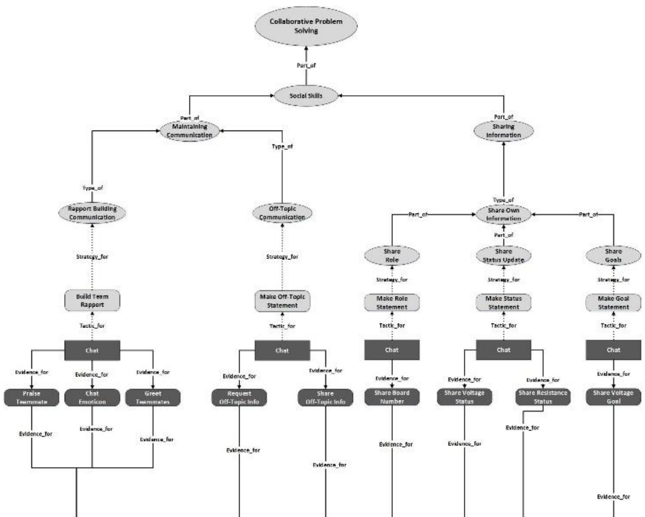

Figure 2. CPS Ontology Fragment

For this particular ontology, the majority of the skills are represented by discourse features associated with team members chatting amongst each other. As there are limited natural language processing tools to identify such low level and abstract features of CPS, qualitative coding was conducted.

2.3 Qualitative Coding

The qualitative coding was conducted on 51,805 rows of log data corresponding to student-generated chats and actions (e.g., resistor changes, calculations) to identify the 23 CPS sub-skills. Specifically, three raters coded each log file event, with each event only receiving one code. To examine inter-rater reliability, a random sample of 20% of the data were triple coded. The inter-rater reliability among the raters was found to be Kappa = .93, indicating substantial agreement [13]. All discrepancies among the coders were discussed to reach consensus for a final code. The remaining data were coded individually by the raters.

During qualitative coding, raters were looking for evidence of one of 23 sub-skills within nine high-level CPS skills under the social and cognitive dimensions of CPS. We next describe each of the sub-skills in turn. In the social dimension, maintaining communication corresponds to content-irrelevant social communication [15,16] and includes three sub-skills, rapport building communication (e.g., greeting teammates), off-topic communication (e.g., talking about homework from another class), and inappropriate communication (e.g., denigrating teammates). Sharing information corresponds to content-relevant information used in the service of solving the problem [22,25] and includes three sub-skills, sharing one's own information (e.g., sharing one's goal voltage), sharing task or resource information (e.g., sharing where the calculator is located in the task interface), and sharing the state of one's understanding (e.g., metacognitive statements such as, "I don't know"). Establishing shared understanding corresponds to communication used to learn the perspective of others and ensure that communication is understood by others [6]. This CPS skill includes two sub-skills, the presentation phase (e.g., requests for information) and the acceptance phase which includes responses indicating comprehension or lack of comprehension of a statement. Negotiating corresponds to communication used to identify conflicts and resolve those that arise [11], and includes three sub-skills, expressing agreement (e.g., "you are right"), expressing disagreement (e.g., "that's not right"), and resolving conflicts.

In the cognitive dimension, exploring and understanding corresponds to actions used to explore the task interface and understand the problem [21] and includes two sub-skills, exploring the environment (e.g., spinning the dial on the DMM) and understanding the problem. Representing and formulating corresponds to communication used to build a mental representation of the problem and formulate hypotheses for how to solve the problem [19,21]. There are two sub-skills for this CPS skill, representing the problem (e.g., "this is a series circuit") and formulating hypotheses (e.g., "I think if everyone has 470 ohms it will be 3.25"). Planning corresponds to communication used to develop a strategy for solving the problem [11, 21], and includes three sub-skills, setting goals (e.g., "We need 6.69 V across our resistors"), managing resources (e.g., "We need to find numbers and decide who does what"), and developing strategies (e.g., "Let's find E first using Kirchhoffs voltage law"). Executing corresponds to communication and actions used in the service of carrying out a plan [21]. This CPS skill includes three sub-skills, enacting strategies (e.g., performing calculations), directing actions (e.g., "Adjust yours to 300 ohms"), and reporting actions (e.g., "I set mine to 120"). Monitoring corresponds to communication and actions in the service of monitoring teammates or progress toward the goal [21,22], and includes two sub-skills, monitoring team organization (e.g., checking on the status of teammates or clicking the Zoom button) and monitoring success (e.g., "We got it" or clicking submit).

3. ANALYSES AND RESULTS

After completing the qualitative coding, the quantitative analyses were conducted in two stages: profile discovery and then validation. For the profile discovery, we performed a hierarchical cluster analysis on the frequencies of each individual's display of the high-level CPS skills using the Ward method [26] as this was an appropriate clustering method given the sample size [17]. We collapsed the 23 sub-skills into the high-level CPS skills for the cluster analysis in order to replicate the process used in previous research [2]. Next, the revealed profiles were compared according to their task performance as identified by number of task levels attempted, performance on the electronics pre-test, and ratings on the self and team CPS Inventory with Kruskal-Wallis tests and Mont Carlo simulations to ensure accurate statistical significance.

3.1 Profile Discovery

We discovered profiles of various types of collaborative problem solvers based on the CPS skills as determined by the competency model (i.e., CPS ontology). Since two of the CPS skills (monitoring and executing) each had both chat and actions as features to determine these skills, we separated them into separate chat and action skills (i.e., monitoring chats, monitoring actions, executing chats, executing actions). Thus, the total number of CPS skills clustered were 11. A hierarchical cluster analysis using the Ward method [26] was conducted on the standardized frequencies of CPS skills displayed for each student. The final number of resulting profiles was determined based on a theoretical interpretation of each of the profiles. Therefore, the profiles were not chosen by fit metrics alone but rather how meaningful these profiles were with respect to social and cognitive psychological research. This was a similar approach to that which was used in our prior work [2]. Although the method was similar, the resulting profiles had some differences. Four profiles emerged with varying sample sizes, which were named Social Loafers, Active Collaborators, Super Socials, and Low Collaborators. In our interpretation of the profiles, we used standardized average frequencies of the CPS skills to discuss patterns across the four profiles.

3.1.1 Social Loafers

The Social Loafers (n = 190) were a group of individuals that displayed below average frequencies of every CPS skill. These individuals did not contribute much to the team's problem solving. Social loafing has a long history in social psychology as a phenomenon where individuals assume that other team members will complete the task and therefore reduce their own effort [14].

3.1.2 Active Collaborators

The Active Collaborators (n = 24) displayed high frequencies on all of the identified CPS skills in the competency model [3] except for monitoring actions (z = -.28). Indeed, these individuals had z values greater than 1 on two of the CPS skills, and greater than 2 on another two CPS skills consistently indicating students being on average, at least an entire standard deviation above the overall mean. Specifically, sharing information (z = 1.44) and establishing shared understanding (z = 1.08), were above the mean. Furthermore, executing chats (z = 2.23) and monitoring chats (z = 2.83) were over a standard deviation above the mean. All other CPS skills had positive standardized values, indicating that these students were generally active in communicating with teammates and helping solve the problem.

3.1.3 Super Socials

The Super Socials (n = 91) showed high frequencies on the social dimension of CPS skills [1], but lower frequencies for the cognitive CPS skills in comparison (except for representing and formulating). Specifically, these individuals showed the highest demonstration of negotiating behaviors in comparison to the other profiles (z = 1.02) and had positive standardized values on all other social skills, though not quite at the level of the Active Collaborators. The only cognitive CPS skills with positive standardized values were communication-based behaviors representing and formulating (z = 1.08), planning (z = .51), executing chats (z =.21), and monitoring chats (z =.09).

3.1.4 Low Collaborators

The Low Collaborators profile (n = 73) consisted of individuals that did not appear to collaborate with their teammates based on the features in the competency model [3]. However, they did show high levels of action-based cognitive behaviors including exploring and understanding (z = .77), monitoring actions (z = 1.21) and executing actions (z = .76). The Low Collaborators had negative standardized values for all other CPS skills. These individuals appeared to work alone without communicating with their teammates which is different from the Social Loafers who simply did not do much work at all.

3.2 Profile Validation

The profiles were validated with both log data performance metrics as well as external measures.

3.2.1 In-Task Performance and Profile Membership

There was a significant relationship between profile membership and the number of task levels attempted, a proxy for performance in the task, (X^2 (3,370) = 7.66, p =.05, *partial* η^2 = .02). The Monte Carlo simulation for significance with 10,000 samples revealed a significance level of p = .05 (lower bound p =.047, upper bound p = .059). Specifically, the mean ranks, where higher values corresponded to more levels attempted, were the lowest for the Social Loafers (175.00) and highest for the Active Collaborators (219.88) which is similar to our previous findings [2]. The mean ranks for the Super Socials and Low Collaborators fell in between the aforementioned profiles (194.87 and 189.57, respectively). These patterns indicate that the Active Collaborators and Super Socials had higher mean ranks on performance than Social Loafers and Low Collaborators.

3.2.2 Pre-Test Performance and Profile Membership

The profiles were compared to the external electronics pre-test, a measure of prior knowledge. The test included 23 items that were summed to create a score for each student participant. Results revealed that there was a significant relationship between profile membership and performance on the electronics test (X^2 (3, 370) = 8.83, p < .05, *partial* η^2 = .02). The Monte Carlo simulation for significance with 10,000 samples revealed a significance level of .027 (lower bound p = .021, upper bound p = .031). The highest mean rank for prior knowledge was for the Active Collaborators (212.73) and the lowest was for the Low Collaborators (172.10). Ranging in the middle, the Super Socials had higher mean ranks than the Social Loafers (209.42 and 175.45, respectively). Post hoc comparisons with a Bonferroni correction revealed a marginally significant difference between Social Loafers and Super Socials (p=.08). No other pairwise comparisons approached statistical significance (all p's > .10).

3.2.3 Post-Task Self-Report and Profile Membership

The profiles were compared to student's ratings of their own CPS behaviors (Self CPS Inventory) and their team's CPS behaviors (Team CPS Inventory).

There was a significant relationship between self-ratings of CPS skills (sum of ratings for Self CPS Inventory) and cluster membership (X^2 (3,349) = 15.57, p < .05, *partial* η^2 = .05). The Monte Carlo simulation with 10,000 samples revealed a significance level of p = .001 (lower bound p = .001, upper bound p = .002). Mean ranks were highest for the Super Socials (210.18) and lowest for the Social Loafers (160.58), with the Active Collaborators having higher mean ranks than the Low Collaborators as expected (184.96 and 162.22, respectively). Post hoc comparisons with a Bonferonni correction revealed a significant difference between Low Collaborators and Super Socials (p < .02) and Social Loafers and Super Socials (p = .001).

There was a significant relationship between ratings on the Team CPS Inventory and profile membership as well (X^2 (3,349) = 9.04, p < .05, *partial* η^2 = .03). Monte Carlo simulation with 10,000 samples revealed a significance of p = .028 (lower bound p = .024, upper bound p = .032). The highest mean rank was for the Super Socials (199.47) and the lowest was for the Social Loafers (161.53), with Active Collaborators having higher mean ranks than Low Collaborators as expected (191.74 and 171.78, respectively). Post hoc comparisons with a Bonferroni correction revealed a significant difference between the Super Socials and Low Collaborators (p = .02).

4. CONCLUSIONS

Overall, we discovered four meaningful profiles of types of collaborative problem solvers: Social Loafers, Active Collaborators, Super Socials, and Low Collaborators. These profiles had significant relationships with in-task performance, electronics prior knowledge, and self-reported CPS capabilities.

The four profiles discovered partially replicate previous findings [2]. Specifically, in our previous study, Social Loafers and Active Collaborators also emerged as profile groups. The Social Loafers could also be called "Free Riders" as these individuals do not contribute much to solving the problem with their teammates. Conversely, the Active Collaborators, which were a small subset of the sample, performed well on all measured aspects of CPS. As expected, Active Collaborators showed better in-task performance than Social Loafers which replicates findings from our prior work [2]. This makes sense as the Active Collaborators displayed high frequencies of CPS behaviors and should therefore have performed well on the task. Social Loafers may have been expecting others to do the work and therefore should not have performed as well on the task.

There were two new profiles that differed but still augmented our previous findings. We attribute these differences to a change in sample size and its diversity. The sample size was nearly three times the size of the previous sample and included students in a wider variety of domains, including electronics, engineering, and physics. The new profiles that emerged in this experiment included the third profile called the Super Socials which does align with other profiles that have been examined. Specifically, in prior work we have found what we termed a high social/low cognitive profile that behaved similarly in displaying high levels of social CPS behaviors and comparatively lower levels of cognitive CPS behaviors. This profile was discovered by an examination of the two dimensions based on means of the CPS features rather than cluster analysis [1]. Beyond this, other work has found a profile designated as "Compensating Collaborators" who had high collaboration actions but performed poorly on problem solving variables [10]. The last profile, the Low Collaborators, also did not emerge in our prior cluster analysis work [2] but could be usefully compared to the Chatty Doers from that work. Similar to the Low Collaborators, the Chatty Doers demonstrated a high level of executing actions, but in contrast, the Chatty Doers did engage in communication with their teammates, though most of the communication was in the maintaining communication category. Interestingly, the Low Collaborators did not seek to engage with their teammates and instead appeared to work alone by engaging in executing and exploratory actions.

In regards to prior knowledge, Active Collaborators and Super Socials demonstrated the first and second highest average scores on the electronics pre-test. It is possible that their higher prior knowledge enabled them to engage in more communication behaviors and problem-solving behaviors (in the case of the Active Collaborators) to contribute to solving the problem. The opposite could be said for the Social Loafers and Low Collaborators, the latter of which had the lowest average pre-test performance. For example, perhaps the Low Collaborators did not want to collaborate with others and preferred to work alone because they were embarrassed of their low levels of content knowledge. On the other hand, perhaps the Low Collaborators already had low content knowledge because of their refusal to work with, and therefore learn from, others on previous tasks. Causality and directionality certainly cannot be determined by these analyses. However, these findings do suggest that testing these hypotheses may provide important insights for CPS researchers.

The Self and Team CPS Inventories required students to rate themselves based on their own metacognitive judgments of their own CPS behaviors as well as their team's CPS behaviors. The Super Socials had the highest ratings for CPS behaviors both for themselves and for their team while the Social Loafers had the lowest ratings on each inventory. These results were expected, though we would also expect high ratings for the Active Collaborators and lower ratings for the Low Collaborators (mean ranks showed such patterns).

All of these findings together suggest that further research should be conducted to explore whether the same kinds of patterns of results emerge with relationships among profiles such as the ones observed in this study, in-task performance, and other ratings. One limitation of this study is that the in-task performance measure includes aspects of the contributions from others while the CPS profile is based on an individual's contributions. Despite the interdependent nature of the electronics task, we are continuing work in developing an alternative in-task performance measure that potentially incorporates only individual contributions. Furthermore, the CPS Inventory relies on self-judgments which can sometimes have biases [8]; however, we did want to incorporate some external measure of CPS behaviors that could be compared to participants' in-task CPS behaviors. Finally, the CPS skills used to develop the profiles included only the higher aggregate level CPS skills, as the sample size was not sufficient to include the lower level coded data.

Overall, we found that this study, which included a larger sample size and new external measures relative to our previous work, partially replicated and informed our previous findings. Our theoretically-grounded data mining approach appears to reveal meaningful profiles on two separate data sets with students completing the same electronics task. We hope that this work will inform future work on ways to incorporate theory and data-driven approaches to make inferences about individuals' CPS capabilities, and contribute to a better understanding of types of collaborative problem solvers, including how certain CPS behaviors relate to various relevant measures.

5. ACKNOWLEDGMENTS

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

6. REFERENCES

[1] Andrews-Todd, J. and Forsyth, C. M. 2020. Exploring social and cognitive dimensions of collaborative problem solving in an open online simulation-based task. *Comp. in Human Beh.*,104,*(2020)*105-759.

[2] Andrews-Todd, J., Forsyth, C. M., Steinberg, J., and Rupp, A. A. 2018. Identifying profiles of collaborative problem solvers in an online electronics environment. In K. E. Boyer & M. Yudelson (Eds.), *Proceedings of the 11th International Conference on Educational Data Mining* (Buffalo, NY July 15-18,2018). EDM '18 International Educational Data Mining Society, 239–245.

[3] Andrews-Todd, J. and Kerr, D. 2019. Application of ontologies for assessing collaborative problem solving skills. *Intern. Journ. of Testing*, 19, 2 (2019), 172–187.

[4] Andrews, J. J. and Rapp, D. N. 2015. Benefits, costs, and challenges of collaboration for learning and memory. *Trans. Issues in Psych. Sci*, 1, 2 (2015), 182-191.

[5] 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 (2013). Educational Testing Service, Princeton, NJ.

[6] Clark, H. H. 1996. *Using Language*. Cambridge University Press, New York, NY.

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

[8] Dunlosky, J. and Metcalfe, J. (2007*). MetaCognition: A Textbook for Cognitive, Educational, Life Span & Applied Psychology.* Sage Publications, Los Angeles, CA.

[9] 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. (*London, U.K. July 4- 7, 2014) EDM '14. International Educational Data Mining Society, 229-232.

[10] 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. *Journ. of Educ. Measur.*, 54, 1, 103–122.

[11] 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. *Assess.t and teaching of 21st century skills.* Springer, New York, NY, 37–56.

[12] 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.

[13] Landis, J. R. and Koch, G. G. 1977. The measurement of observer agreement for categorical data. *Biometrics*, 33,1 (1977), 159–174.

[14] Latané, B., Williams, K., and Harkins, S. 1979. Many hands make light the work: The causes and consequences of social loafing. *Journ. of Pers. and Soc. Psycho.* 37, 6 (1979), 822-832.

[15] Lipponen, L. (2000). Towards knowledge building: From facts to explanations in primary students' computer mediated discourse. *Learn. Environ. Res*, *3*, 2 (2000), 179–199.

[16] 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,(2015) 344–359.

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

[18] Meier, A., Spada, H., and Rummel, N. 2007. A rating scheme for assessing the quality of computer-supported collaboration processes. *Inter,l Journ. of Computer-Supported Collab. Learning*, 2, 1 (2007), 63–86.

[19] Mayer, R. E. and Wittrock, M. C. 1996. Problem-solving transfer. In D. C. Berliner & R. C. Calfee (Eds.), *Handbook of educational psychology* (1996) Lawrence Erlbaum, 47–62.

[20] Morgan, B. B., Salas, E., and Glickman, A. S. 1993. An analysis of team evolution and maturation. *Journal of Gen. Psych.*, 120, 3 (1993), 277–291.

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

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

[23] 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 (2003), 361–373

[24] Steinberg, J., Andrews-Todd, J., Forsyth, C., Chamberlain, J., Horwitz, P., Koon, A., Rupp, A. A., and McCulla, L. 2019. *The Development of a Content Assessment of Basic Electronics Knowledge*. Paper presented at 50th Annual Conference of the Northeastern Educational Research Association, Trumbull, CT. (October, 2019).

[25] van Boxtel, C., van der Linden, J., and Kanselaar, G. 2000. Collaborative learning tasks and the elaboration of conceptual knowledge. *Learn. and Instruc*, 10, 4 (2000), 311–330.

[26] Ward, Jr., J. H. 1963. Hierarchical grouping to optimize an objective function. *Journ, of the American Stat. Assoc.*, 58, 301 (1963), 236-244. DOI: 10.2307/2282967