Computer-based Assessment of Collaborative Problem Solving Skills: Human-to-Agent versus Human-to-Human

Yigal Rosen, Ph.D.

Pearson, USA

yigal.rosen@pearson.com

and

Maryam Tager

Pearson, USA

maryam.tager@pearson.com

Abstract

In order to understand how students perform on collaborative problem solving (CPS) computer-based assessment, it is necessary to examine empirically the multi-faceted performance that may be distributed across collaboration methods. The aim of this study was to explore possible differences in student performance in human-to-agent (H-A), compared to human-to-human (H-H) CPS assessment tasks. One hundred seventy nine 14 years-old students from the United States, Singapore and Israel participated in the study. Students in both H-H and H-A modes were able to collaborate and communicate by using identical methods and resources. However, while in the H-A mode, students collaborated with a simulated computer-driven partner and in the H-H mode students collaborated with another student to solve a problem. Overall, the findings showed that CPS with a computer agent involved significantly higher levels of shared understanding, progress monitoring, and feedback. However, no significant difference was found in a student's ability to solve the problem or in student motivation with a computer agent or a human partner. This study is among the first of its kind to investigate systematically the effect of collaborative problem solving in H-A and H-H standardized assessment settings.

KEYWORDS: Collaborative problem solving, computer-based assessment, groups

Introduction

Collaborative problem solving (CPS) is a critical competency for college and career readiness. Students emerging from schools into the workforce and public life will be expected to have CPS skills as well as the ability to perform that collaboration in various group compositions and environments (Griffin, Care, & McGaw, 2012; OECD, 2013; O'Neil, & Chuang, 2008; Rosen, & Rimor, 2012). Recent curriculum and instruction reforms have focused to a greater extent on teaching and learning CPS (National Research Council, 2011). However, structuring standardized computer-based assessment of CPS skills, specifically for large-scale assessment programs, is challenging. In a standardized assessment situation, a student should be matched with various types of group members that will represent different CPS skills and contexts. In addition, the discourse between the group members should be manageable and predictable. The two major questions thus are: *Can partners for CPS be simulated but still maintain authentic human aspects of collaboration? And, how can manageable and predictable group discourse spaces be created within the assessment?* This paper addresses these challenges by introducing a new methodology for scalable computer-based assessment of CPS, proving findings from an empirical pilot study conducted in three counties, as well as discussing implication of the findings on further research and development.

Defining Collaborative Problem Solving

Currently, the terms "collaborative problem solving", "cooperative work" and "group work" are used interchangeably in the education research literature to mean similar constructs. Collaborative problem solving thus refers to problem-solving activities that involve collaboration among a group of individuals (O'Neil, Chuang, & Baker, 2010). CPS is a conjoint construct consisting of collaboration, or: "coordinated, synchronous activity that is the result of a continued attempt to construct and maintain a shared conception of a problem" (Roschelle, & Teasley, 1995, p. 70), and problem solving, or: "cognitive processing directed at achieving a goal when no solution method is obvious to the problem solver" (Mayer, & Wittrock, 1996). According to Griffin, Care, and McGaw (2012), CPS refers to the ability to recognize the points of view of other persons in a group; to contribute knowledge, experience, and expertise in a constructive way; to identify the need for contributions and how to manage them; to recognize structure and procedure involved in resolving a problem; and as a member of the group, to build and develop group knowledge and understanding. CPS is one of the two major areas that the Organisation for Economic Co-operation and Development (OECD) selected in 2015 for primary development in Programme for International Student Assessment (PISA). In PISA 2015, CPS competency is 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" (OECD, 2013). An agent could be considered either a human agent or a computer agent that interacts with the student. The competency is assessed by evaluating how well the individual collaborates with agents during the problem-solving process. This includes establishing and maintaining shared understanding, taking appropriate actions to solve the problem, and establishing and maintaining group organization.

In our research, an operational definition of CPS refers to "*the capacity of an individual to effectively engage in a group process whereby two or more agents attempt to solve a problem by sharing knowledge and understanding, organizing the group work and monitoring the progress, taking actions to solve the problem, and providing constructive feedback to group members.*"

Assessing Collaborative Problem Solving Skills

Student performance in CPS can be assessed through a number of different methods. These include measures of the quality of the solutions and the objects generated during the collaboration (Avouris, Dimitracopoulou, & Komis, 2003); analyses of log files, intermediate results, paths to the solutions (Adejumo et al., 2008), team processes and structure of interactions (O'Neil, Chung, & Brown, 1997); and quality and type of collaborative communication (Foltz, & Martin, 2008; Graesser, Jeon, & Dufty, 2008). There are distinct tradeoffs between the large amount of information that can be collected in a collaborative activity and what can be measured. For example, while the content of spoken communications is quite rich, analyses by human markers can be quite time consuming and difficult to automate. Nevertheless, much of the problem-solving process data as well other communication information (turn taking, amount of information conveyed) can be analyzed by automatic methods.

To ensure valid measurement on the individual level, each student should be paired with the same number of other partners displaying the same range of CPS characteristics. This way each individual student will be situated fairly similarly to be able to show his or her proficiency in CPS. A key consideration in CPS assessment is the development of situations where collaboration is critical to performing successfully on the task. Such situations require interdependency between the students where information must be shared. For example, dynamic problem situations can be developed where each team member has a piece of information and only together can they solve the problem (called hidden-profile or jigsaw problems). Similarly, ill-defined tasks can be developed using such tasks as group bargaining where there are limited resources but a group must converge on a solution that satisfies the needs of different stakeholders. Finally, information between participants may also be conflicting, requiring sharing of the information and then resolution in order to determine what information best solves the problem. CPS tasks may include one or more of these situations, while the common factors in all these collaborative tasks are handling discord, disequilibrium, and group think. Usually, a group member cannot complete the task without taking actions to ensure that a shared understanding is established. Thus, a key element in CPS tasks is interdependency between the group members so that grounding is both required and observable.

Human-to-Human and Human-to-Agent Approach in CPS Assessment

Collaboration can take many forms, ranging from two individuals to large teams with predefined roles. The Human-to-Human (H-H) approach provides an authentic human-human interaction which is a highly familiar situation for students. Students may be more engaged and motivated to collaborate with their peers. Additionally, the H-H situation is closer to the CPS situations students will encounter in their personal, educational, professional and civic activities. However, pairing can be problematic because of individual differences that can significantly affect the CPS process and its outcome. Therefore, the H-H assessment approach of CPS may not provide enough opportunity to cover variations in group composition, diversity of perspectives and different team member characteristics in controlled manners, which are all essential for assessment on an individual level. Simulated team members for collaboration with a preprogrammed profile, actions and communication would potentially provide the coverage of the full range of collaboration skills with sufficient control. In the Human-to-Agent (H-A) approach, CPS skills are measured by pairing each individual student with a computer agent or agents that can be programmed to act as team members with varying characteristics relevant to different CPS situations. Group processes are often different depending on the task and could even be competitive. Use of computer agents provides an essential component of non-competitiveness to the CPS situation, as it is experienced by a student. Additionally, if the time-on-task is limited, taking the time to explain to each other may lower group productivity. As a result of these perceived constraints, a student collaborating in H-H mode may limit significantly the extent to which CPS dimensions, such as shared understanding, are externalized through communication with the partner. The agents in H-A communication can be developed with a full range of capabilities, such as text-to-speech, facial actions, and optionally rudimentary gestures. However, CPS in H-A settings deviate from natural human communication delivery and can cause distraction and sometimes irritation. The dynamics of H-H interaction (timing, conditional branching) cannot be perfectly captured with agents. For example, human collaborators can propose unusual, exceptional solutions; the characteristic of such a process is that it cannot be included in a system following an algorithm, such as H-A interaction.

Research shows that computer agents can be successfully used for tutoring, collaborative learning, co-construction of knowledge, and CPS (e.g., Biswas et al., 2010; Millis et al., 2011). A computer agent can be capable of generating goals, performing actions, communicating messages, sensing its environment, adapting to changing environments, and learning.

In summary, CPS assessment must take into account the types of technology, tasks and assessment contexts in which it will be applied. The assessment will need to consider the kinds of constructs that can be reliably measured and also provide valid inferences about the collaborative skills being measured. Computer-based assessment of CPS involves the need for advancements in educational assessment methodologies and technology. The paper addresses these challenges by studying student CPS performance in two modes of CPS assessment.

Research Questions

The study addressed empirically the following primary question regarding students' CPS performance in H-A, compared to H-H CPS settings:

What are the differences in student CPS performance between H-A and H-H mode of assessment, as reflected in shared understanding, problem solving, progress monitoring and proving feedback measures?

In order to better understand the dimensionality of CPS measures and possible factors that differentiate student performance in H-A and H-H settings, the following research questions were examined:

What are the relationships between the CPS measures of shared understanding, problem solving, progress monitoring, and proving feedback measures in H-A and H-H settings?

What are the differences in student motivation while collaborating with a computer agent or a human partner on CPS assessment tasks?

What are the differences in student CPS performance between H-A and H-H modes of assessment, as reflected in time-on-task, and number of attempts to solve the problem?

Method

Study participants included 179 students age 14, from the United States, Singapore and Israel. The results presented in the current article came from a larger study in which students from six countries were recruited to participate in a 21st Century Skills Assessment project investigating the innovative ways to develop computer-based assessment of critical-thinking, and CPS. The researchers collected data between November 2012 and January 2013. Recruitment of participating schools was achieved through collaboration with local educational organizations based on the following criteria: (a) the school is public, (b) the school is actively involved in various 21st Century Skills projects, (c) the population is 14 years-old students proficient in English, and (c) there is sufficient technology infrastructure (e.g. computers per student, high-speed Internet). In all, 136 students participated in the H-A group and 43 participated in the H-H group (43 additional students participated in the H-H setting, acting as 'collaborators' for the major H-H group). Specifically in H-H assessment mode, students were randomly assigned into pairs to work on the CPS task. Because the H-H approach required pairs of students working together in a synchronized manner, the number of pairs was limited. This is due to the characteristics of technology infrastructures in participating schools.

Of the total students who participated, 88 were boys (49.2%) and 91 were girls (50.8%). No significant differences were found in Grade Point Average (GPA), English Language Arts (ELA), and Math average scores between participants in H-A and H-H mode within the countries. This similar student background allowed comparability of student results in CPS assessment task between the two modes of collaboration.

Collaborative Problem Solving Assessment

In this CPS computer-based assessment task, the student was asked to collaborate with a partner (computer-driven agent or a classmate) to find the optimal conditions for an animal at the zoo. The student was able to select different types of food, life environments, and extra features, while both partners were able to see the selections made and communicate through a phrase-chat (selections from predefined 4-5 options). An animal's life expectancy under the given conditions was presented after each trial of the conditions. The student and the partner were prompted to discuss how to reach better conditions for an animal at the beginning of the task. By the end of the task, the student was asked to rate the partner (1-3 stars) and provide written feedback on the partner's performance. It should be noted that due to the centrality of the collaboration dimension in CPS as it was defined in this study, the difficulty level of the problem was relatively low and served primarily as a platform for the overall assessment of CPS skills. Additionally, due to the exploratory nature of the study, the students were not limited either in a number of attempts to reach optimal solution or in the time-on-task. However, the task was programmed in such a way that at least two attempts for problem solving and at least one communication act with a partner were required to be able to complete the assessment task.

The task was checked with ten teachers from the three participating countries to ensure that students would be able to work on the task, that the task could differentiate between high and low levels of CPS ability, and that the task was free of cultural biases. Interviews were conducted with eight students representing the target population to validate various CPS actions and communication programmed for the computer agent and to establish automatic scoring of student responses.

CPS scores for the assessment task consisted of shared understanding (40 points), problem solving (26 points), monitoring progress (26 points), and providing feedback (8 points). Both in H-H and H-A settings, student scores in the first three CPS dimensions were generated automatically based on a predefined programmed sequence of possible optimal actions and communication that was embedded into the assessment task. Scoring of student feedback was provided independently by two teachers from participating schools in the United States. Spelling and grammar issues did not affect student score. Inter-coded agreement of feedback scoring was 92%.

Questionnaire

The questionnaire included four items to assess the extent to which students were motivated to work on the task. Participants reported the degree of their agreement with each item on a four-point Likert scale (1 = strongly disagree, 4 = strongly agree). The items were adopted from motivation questionnaires used in previous studies, and included: "I felt interested in the task"; "The task was fun"; "The task was attractive"; "I continued to work on this task out of curiosity" (Rosen, 2009; Rosen, Beck-Hill, 2012). The reliability (internal consistency) of the questionnaire was .85. Students were also asked to indicate the background information, including gender, GPA, and Math and ELA average score, as measured by school assessments.

Results

All results are presented on an aggregative level beyond the countries, since no interaction with student-related country was found. First, the results of student performance in a CPS assessment are presented to determine whether there is a difference in student CPS score as a function of collaborating with a computer agent versus a classmate. Next, student motivation results are presented to indicate possible differences in H-A and H-H modes. Last, time-on-task and number of attempts to solve the problem in both modes of collaboration are demonstrated.

Comparing Student CPS Performance in H-H and H-A settings

In order to explore possible differences in students' CPS scores analysis of variance was performed. First, MANOVA results showed significant difference between H-H and H-A groups (Wilks' Lambda=.904, $F(df=4,174)=4.6$, $p<.01$). Hence, we proceed to perform t-tests. The results indicated that students who collaborated with a computer agent showed significantly higher level of performance in establishing and maintaining shared understanding (ES=.4, $t(df=177)=2.5$, $p<.05$), monitoring progress of solving the problem (ES=.6, $t(df=177)=4.0$, $p<.01$), and in the quality of the feedback (ES=.5, $t(df=177)=3.2$, $p<.01$). The findings showed non-significant difference in the ability to solve the problem in the H-A and H-H mode of collaboration (ES=-.3, $t(df=177)=-1.9$, $p=.06$).

Relationships between CPS Scores

To better understand the relationship between the CPS scores, analysis of intercorrelations between the variables was conducted. The analysis was conducted separately in H-H and H-A conditions because of the possible differences in the intercorrelations across different collaboration settings. The H-H findings showed significantly positive relationships between student shared understanding score and the ability to monitor progress ($r=.43$, $p < .01$), while student problem-solving score was negatively correlated with shared understanding ($r=-.34$, $p <$

.05). The H-A setting revealed a similar pattern, mainly in relationship between shared understanding and monitoring progress score ($r=.57$, $p < .01$). In addition, shared understanding score was found in significantly positive correlation with student ability of providing constructive feedback to the partner ($r=.48$, $p < .01$). No correlation was found between student problem solving score and other CPS measures.

Student Motivation

In attempting to determine possible differences in student motivation of being engaged in CPS with a computer agent versus a classmate, data on student motivation was analyzed. The result demonstrated that it is a matter of indifference in student's motivation whether collaborating with a computer agent or a classmate (M=3.1, SD=.7 in H-A mode, compared to M=3.1, SD=.4 in H-H mode; ES=.1, $t(df=177)=.5$, $p=.64$).

Attempts to Solve a Problem and Time-on-Task

In order to examine possible differences in the number of attempts for problem-solving as well as time-on-task, a comparison of these measures was conducted between H-A and H-H modes of collaboration. In practice, the average number of attempts for problem solving in H-A mode was 8.4 (SD=7.3), compared to 6.1 (SD=5.7) in a H-H mode (ES=.3, $t(df=177)=2.1$, $p<.05$). No significant difference was found in time-on-task (ES=-1.9, $t(df=177)=-1.6$, $p=.11$). On average, time-on-task in H-A mode was 7.9 minutes (SD=3.6), while student in the H-H mode spent 1.1 more minutes on a task (M=9.0, SD=4.5).

Discussion

The goal of this study was to explore differences in student CPS performance in H-A and H-H modes. Students in each of these modes were exposed to identical assessment tasks and were able to collaborate and communicate by using identical methods and resources. However, while in the H-A mode students collaborated with a simulated computer-driven partner, and in the H-H mode students collaborated with another student to solve a problem. The findings showed that students assessed in H-A mode outperformed their peers in H-H mode in their collaborative skills. CPS with a computer agent involved significantly higher levels of shared understanding, progress monitoring, and feedback. The results suggest that the space of collaboration in H-A settings can be extremely large even when there are a limited number of fixed actions or discourse moves at each point in a conversation. Although students in both H-H and H-A modes were able to collaborate and communicate by using identical methods and resources, full comparability was not expected. This is due to the fact that each student in H-H mode represented a specific set of CPS skills, while in the H-A mode each individual student collaborated with a computer agent with a predetermined large spectrum of CPS skills. Differences across H-H groups could be affected by a given performance of the collaborator. Additionally, because of the relatively low difficulty of the problem that was represented by the CPS task, and much larger emphasis on collaboration, students in H-A were faced with more opportunities to show their collaboration skills. Research shows that in H-H CPS settings there is a tendency to avoid disagreements in order to achieve a rapid consensus on how to solve a problem (e.g., Rosen, & Rimor, 2012). It is possible that some students that acted as collaborators in H-H settings did not involve themselves in disagreements, questioning, alternative interpretations of results and other possible resources for sharing understanding, monitoring progress, and providing feedback that can be performed by the leader student. This was not the case with a computer agent. The agent was programmed to partially disagree with the student, occasionally misinterpret the results, or propose misleading strategies.

One major possible implication of CPS score difference in collaboration measures between the H-A and H-H modes is that assessments delivered in multiple modes may differ in score meaning and impact. Each mode of CPS assessment can be differently effective for different educational purposes. For example, a formative assessment program which has adopted rich training on the communication and collaboration construct for its teachers may consider the H-H approach for CPS assessment as a more powerful tool to inform teaching and learning, while H-A may be implemented as a formative scalable tool across a large district or in standardized summative settings. Non-availability of students with a certain CPS level in a class may limit the fulfilment of assessment needs, but technology with computer agents can fill the gaps. In many cases, using simulated computer agents instead of relying on peers is not merely a replacement with limitations, but an enhancement of the capabilities that makes independent assessment possible. Furthermore, a phrase-chat used in this study can be replaced by an open-chat in cases where automated scoring of student responses is not needed.

In contrast, we found that it is a matter of indifference in student ability to solve the problem with a computer agent or a human partner, although on average students in H-A mode applied more attempts to solve the problem, compared to the H-H mode. Student performance studied here was in the context of well-structured problem-solving, while primarily targeting collaborative dimensions of CPS. The problem-solving performance in this task was strongly influenced by the ability of the students to apply a vary-one-thing-at-a-time strategy (Vollmeyer, & Rheinberg, 1999), which is also known as control of variables strategy (Chen, & Klahr, 1999). This is a method for creating experiments in which a single contrast is made between experimental conditions. While the computer agent was programmed to suggest this strategy to each participant in a standardized way (before the second submission), there was no control over the suggestions made by the human partner. However, as shown in this study, participants in H-A mode do not outperform participants in H-H mode in their problem-solving score, while the major difference between the students' performance in H-H and H-A settings are the collaboration-related skills. In fact, correlation analysis of students' CPS scores indicated that while some of the collaborative dimensions were positively intercorrelated in both H-H and H-A settings, there was no direct connection between the collaborative effort in the task and the ability to solve the problem (i.e., no correlation in H-A and low-level negative correlation in H-H settings). Interdependency is a central property of tasks that are desired for assessing collaborative problem solving, as opposed to a collection of independent individual problem solvers. A task has higher interdependency to the extent that student A cannot solve a problem without actions of student B. Although, interdependency between the group members was required and observable in the studied CPS task, the collaboration in both settings was characterized by asymmetry of roles. A "leader" student in the H-H setting and the student in the H-A setting were in charge of selecting the variables and submitting the solutions in addition to the ability to communicate with the partner. According to Dillenbourg (1999), asymmetry of roles in collaborative tasks could affect each team member's performance. Thus, a possible explanation for these results is the asymmetry in roles between the "leader" student and the "collaborator" in the H-H setting and the student and the computer agent in the H-A setting. In a more controlled setting (i.e., H-A) the asymmetrical nature of collaboration was associated with no relationship to the quality of collaborative skills that were observed during the task. While in the H-H setting, in which the human "collaborator" was functioning with no system control over the suggestions that he or she made, the asymmetry in roles was associated with no relationship to the quality of collaborative skills that were observed during the task. A major factor that contributes to the success of CPS and differentiates it from individual problem solving is the role of communication among team members (Fiore, & Schooler, 2004; Dillenbourg, & Traum, 2006; Fiore et al., 2010). Communication is essential for organizing the team, establishing a common ground and vision, assigning tasks, tracking progress, building consensus, managing conflict, and a host of other activities in CPS.

Concerning the level of motivation and time-on-task in collaborating with a computer agent or a human partner on CPS assessment task, we found no evidence for differences between the two modes. In other words, students felt motivated and efficient in collaborative work with computer agents just at the same level as when collaborating with their peers. Previous research found that examinee motivation tended to predict test performance among students in situations in which the tests had low or no stakes for the examinees (Wise, & DeMars, 2005). To the degree to which students do not give full effort to an assessment test, the resulting test scores will tend to underestimate their levels of proficiency (Eklöf, 2006; Wise & DeMars, 2005). We believe that two major factors in computer agent implementation contributed to student motivation in CPS assessment tasks. On the one hand, the student and the agent shared the responsibility to collaborate in order to solve the problem. A computer agent was capable to generate suggestions to solve the problem (e.g., "Let's change one condition per trial") and communicate with the student in a contextual and realistic manner. On the other hand, a shared representation of the problem-solving space was implemented to provide a concrete representation of the problem state (i.e., life expectancy) and the selections made (e.g., selection of the conditions).

In summary, the results of this study suggest that by using computer agents in a CPS task the students were able to show their collaborative skills at least at the level of that of their peers who collaborated with human partners. However, as discussed in this article, each mode of collaboration involves limitations and challenges. Further research is needed in order to establish comprehensive validity evidence and generalization of findings both in H-A and H-H CPS settings.

References

Adejumo, G., Duimering, R. P. Zhong, Z. (2008). A balance theory approach to group problem solving, *Social Networks, 30*(1), 83-99.

Avouris N., Dimitracopoulou A., & Komis V. (2003). On analysis of collaborative problem solving: An object-oriented approach. *Computers in Human Behaviour, 19*(2), 147-167.

Biswas, G., Jeong, H., Kinnebrew, J. S., Sulcer, B., & Roscoe, A. R. (2010). Measuring self-regulated learning skills through social interactions in a teachable agent environment. *Research and Practice in Technology-Enhanced Learning, 5*(2), 123-152.

Dillenbourg, P. (Ed.). (1999). *Collaborative learning: Cognitive and computational approaches*. Amsterdam, NL: Pergamon, Elsevier Science.

Dillenbourg, P., & Traum, D. (2006). Sharing solutions: Persistence and grounding in multi-modal collaborative problem solving. *The Journal of the Learning Sciences, 15*(1), 121-151.

Eklöf, H. (2006). Development and validation of scores from an instrument measuring student test-taking motivation. *Educational and Psychological Measurement, 66*(4), 643-656.

Fiore, S., Rosen, M., Smith-Jentsch, K., Salas, E., Letsky, M., & Warner, N. (2010). Toward an understanding of macrocognition in teams: Predicting process in complex collaborative contexts. *The Journal of the Human Factors and Ergonomics Society*, 53(2), 203-224.

Fiore, S., & Schooler, J. W. (2004). Process mapping and shared cognition: Teamwork and the development of shared problem models. In E. Salas, & S. M. Fiore (Eds.), *Team cognition: Understanding the factors that drive process and performance* (pp. 133–152). Washington DC: American Psychological Association.

Foltz, P. W. & Martin, M. J. (2008). Automated communication analysis of teams. In E. Salas, G. F. Goodwin, & S. Burke (Eds.) Team effectiveness in complex organizations and systems: Cross-disciplinary perspectives and approaches (pp. 411-431). New York: Routledge.

Griffin, P., Care, E., & McGaw, B. (2012). The changing role of education and schools. In P. Griffin, B. McGaw, & E. Care (Eds.), *Assessment and teaching 21st century skills* (pp. 1-15). Heidelberg: Springer.

Mayer, R. E., & Wittrock, M. C. (1996). Problem-solving transfer. In D. C. Berliner & R. C. Calfee (Eds.), Handbook of educational psychology (pp. 47-62). New York: Macmillan Library Reference USA, Simon & Schuster Macmillan.

Millis, K., Forsyth, C., Butler, H., Wallace, P., Graesser, A. C., & Halpern, D. (2011). Operation ARIES! A serious game for teaching scientific inquiry. In M. Ma, A. Oikonomou, & J. Lakhmi (Eds.), *Serious games and edutainment applications* (pp. 169-195). London: Springer-Verlag.

National Research Council (2011). *Assessing 21st Century Skills. National* Academies Press, Washington, DC.

OECD (2013). *PISA 2015 Collaborative Problem Solving Framework*. OECD Publishing.

O'Neil, H. F., Jr., & Chuang, S. H. (2008). Measuring collaborative problem solving in low-stakes tests. In E. L. Baker, J. Dickieson, W. Wulfeck, & H. F. O'Neil (Eds.), Assessment of problem solving using simulations (pp. 177-199). Mahwah, NJ: Lawrence Erlbaum Associates.

O'Neil, H. F., Jr., Chuang, S. H., & Baker, E. L. (2010), Computer-based feedback for computer-based collaborative problem solving. In D. Ifenthaler, P. Pirnay-Dummer, & N. M. Seel (Eds.), *Computer-based Diagnostics and Systematic Analysis of Knowledge* (pp. 261-279). New York: Springer-Verlag.

Roschelle, J. & Teasley, S. D. (1995). The construction of shared knowledge in collaborative problem-solving. In C. E. O'Malley (Ed.), *Computer-supported collaborative learning* (pp. 69–97). Berlin: Springer-Verlag.

Rosen, Y. (2009). Effects of animation learning environment on knowledge transfer and learning motivation. *Journal of Educational Computing Research, 40*(4), 439-455.

Rosen, Y., & Beck-Hill, D. (2012). Intertwining digital content and one-to-one laptop learning environment. *Journal of Research on Technology in Education, 44*(3), 223-239.

Rosen, Y., & Rimor, R. (2012). Teaching and assessing problem solving in online collaborative environment. In R. Hartshorne, T. Heafner, & T. Petty (Eds.), *Teacher education programs and online learning tools: Innovations in teacher preparation* (pp. 82-97). Hershey, PA: Information Science Reference, IGI Global.

Wise, S. L., & DeMars, C. E. (2005). Low examinee effort in low-stakes assessment: Problems and potential solutions. *Educational Assessment*, *10*(1), 1-17.