Yike Ying[1]
Rüdiger Tiemann[1]

[1]Humboldt-Universität zu Berlin

Development and evaluation of an assessment tool for collaborative problem-solving skills in chemistry (CPS-C)

Theoretical background

Collaborative problem solving (CPS) is an important component of 21st century skills and is considered a vital competency in today's connected world (Griffin & Care, 2014; Hendarwati et al., 2021). It generally refers to a situation where two or more people pool their knowledge and skills to solve complex problems without predefined solutions (Sun et al., 2022). Particularly in the field of education, CPS is considered a complex, multidimensional, multi-level skill that typically places learning in the context of solving real-world, unstructured problems that require students to coordinate social, cognitive, metacognitive, and behavioral aspects over time (Graesser et al., 2018; Hmelo-Silver & DeSimone, 2013). In addition, it encourages students to create group knowledge and build responsibility, self-regulation, and collaboration skills for learning. CPS has been used extensively in K-12 education, higher education, and informal learning to improve the quality of student learning. The main studies focusing on the assessment of CPS have been in the Assessment and Teaching of 21st Century Skills project (ATC21S) and in the Program for International Student Assessment (PISA). The approach taken by ATC21S identified two domains of CPS, social and cognitive (Griffin & Care, 2014), while the PISA framework includes four problem-solving processes and three collaborative competences (OECD, 2013). Furthermore, the studies in ATC21S involved human-to-human (H-H) collaborative interaction, whereas the PISA assessment involved human-to-agent (H-A) interaction. Unlike the H-H approach in ATC21S, the H-A approach in the PISA assessment were controllable and standardized, placing students in a variety of different collaborative situations, as well as controlling test time (Wang, 2016). What's more, students interacted with one or more computer agents, all questions were derived from real-life problem scenarios, and the computer agents gave instructions and feedback based on their answers. For chemistry, which is an experiment-based subject, experimentation is an important means and method for students to apply their theoretical knowledge in practice through collaborative group chemistry experiments.

Research aims

So far, the research investigating the contributions and effects of conversational agents in collaborative situations is still scarce (Kuo et al., 2019); only few studies have examined online CPS in science (Slotta & Linn, 2000), and even fewer in chemistry, which had led to a demand for CPS evaluation tools. In addition, student variables such as gender, age, interests, attitudes, stress, etc. may have an impact on CPS performance. However, studies exploring factors that predict CPS competence have been almost exclusively qualitative, with insufficient quantitative evidence (Tang et al., 2021). For the above reasons, the research aims of this project are to:

- Develop and evaluation of a standardized CPS assessment tool by H-A method in chemistry (CPS-C).

- Discover the influence of interest, motivation, cognitive ability, and other factors on student performance in CPS in chemistry.

Methods and design

In the PISA framework of CPS, the three collaborative skills and the four problem-solving skills comprise a matrix of 12 CPS skills (see OECD, 2017, p. 12 for details). We developed in total 60 multiple-choice items based on these 12 skills to record students' responses. These tasks use LimeSurvey (Version 5.3.13; Limesurvey GmbH, 2012) as a technology-based tool and contain three main topics: Coca-Cola titration (1 agent, 19 items), fruit-battery (2 agents, 21 items) and Soap-making (3 agents, 20 items). To ensure uniformity across participants and actual events, the agent's characteristics are pre-programmed. In a preliminary qualitative study, N=52 (34 male/18 female) grade 11 students in Chinese general high schools were selected for testing and N=10 were selected for interviews at the end of the test in order to get their feedback. Following that, N=292 (205 male/87 female) students participated in the main study. For both the preliminary and main study, additional assessments were administered alongside the CPS-C to explore the factors influencing collaborate problem-solving performance: basic cognitive abilities (KFT-10; Heller & Perleth, 2000), interest and motivation (Rost, 2021), mental load (Krell, 2015), stress (Minkley et al., 2018) and prior knowledge.

To estimate the student’s CPS-C performance, each item only measured one specific dimension and their raw scores were scaled with the help of MIRT models in R (Version 4.2.1; R Core Team, 2022) and the generalized partial credit model (GPCM; Muraki, 1992) was selected to rate the item scores. After verification of the CPS-C test, we used multiple linear regression analysis to find which variable has the great effect on CPS-C.

Results

This study used the EAP/PV reliability coefficient to examine students' accurate estimates of item fit in the collaboration competencies and problem-solving processes, and it assessed the fit of the data to the applied MIRT model at the item level using weighted mean square (infit) (see Table 1&2). Overall, the results showed acceptable internal consistency. The results of the reliability coefficients in the three collaboration competencies were respectively 0.704, 0.692, and 0.635. In terms of the four problem solving processes, the reliability coefficients were respectively 0.718, 0.619, 0.611 and 0.610. Moreover, all of the items showed that the expected value of the infit was between 0.8 and 1.1, indicating a good fit with the CPS-C model. In addition, for student’s performance in CPS-C, the overall mean latent ability was around 1.2, suggesting that the students were willing to share perspectives, negotiate with team members, and resolve the problem they encountered. However, students performed better in the Monitoring and reflecting steps, with a latent ability of 1.437. According to the results, the Chinese students were more concerned with the results of the collaborative process. The results of the correlation analysis showed that a significant negative correlation between stress ($r=-0.169$, $p<0.01$) and CPS-C and a significant correlation between CPS-C and the KFT ($r=0.469$, $p<0.01$) and prior-knowledge ($r=0.367$, $p<0.01$) scores. In order to find which variables had the greatest effect on CPS, a multiple linear regression analysis was performed. The regression model coefficient ($R=0.527$, $R2=0.277$; $F(3, 289)=48.177$, $P<0.001$) indicated that there is a linear correlation between the dependent and independent variables and that the regression model is statistically significant. The effects of the KFT (B=0.722, p<0.001, VIF=1.112) and prior knowledge (B=1.636, p<0.001, VIF=1.159) scores on CPS-C were statistically significant, while Stress (B=-0.317,p=.0.345>0.05, VIF=1.076) was not.

Table 1. Reliability, item-fit and ability mean of collaboration competencies

Competencies of collaboration	**Number of items**	**Reliability coefficients (EAP/PV)**	**Weighted fit MNSQ (infit range)**	**Ability mean (SD)**
Establishing and maintaining a shared understanding	20	0.704	0.878-0.998	1.224 (0.581)
Taking the appropriate action to solve the problem	20	0.692	0.932-1.031	1.265 (0.616)
Establishing and maintaining team organization	20	0.635	0.926-1.002	1.197 (0.646)

Table 2. Reliability, item-fit and ability mean of problem solving processes

Competencies of problem solving	**Number of items**	**Reliability coefficients (EAP/PV)**	**Weighted fit MNSQ (infit range)**	**Ability mean (SD)**
Exploring and understanding	15	0.718	0.907–0.972	1.201 (0.609)
Representing and formulating	14	0.619	0.907–0.995	1.200 (0.633)
Planning and executing	16	0.611	0.932–1.073	1.222 (0.634)
Monitoring and reflecting	15	0.610	0.869–0.991	1.437 (0.645)

Conclusion and discussion

The results indicate that the CPS-C assessment tool was practical and capable of tracking changes in student performance. Although the reliability is unsatisfactory but within acceptable bounds, the goodness of fit assessment of MIRT yields satisfactory results, which may be partially attributed to the small sample size. For student performance in CPS-C, students have intermediate levels of CPS-C competency and are able to convey a common understanding of the issue and required knowledge to solve problems. Students performed better in Monitoring and Reflecting than they did in the other CPS-C dimensions, indicating that the Chinese students appear to pay attention to the results of experiments and provide timely feedback and moderation. Another reason for this result could be the unbalanced total score for each dimension, with some questions being dichotomous and others being polytomous. Moreover, an exploration of variables related to CPS performance found that prior knowledge and cognitive ability had a significant effect on students' CPS performance.

Literatur

Graesser, A. C., Fiore, S. M., Greiff, S., Andrews-Todd, J., Foltz, P. W., & Hesse, F. W. (2018). Advancing the science of collaborative problem solving. Psychological Science in the Public Interest, 19(2), 59-92.

Griffin, P., & Care, E. (2014). Assessment and teaching of 21st century skills: Methods and approach. Springer.

Hendarwati, E., Nurlaela, L., Bachri, B., & Sa'ida, N. (2021). Collaborative Problem based learning integrated with online learning. International Journal of Emerging Technologies in Learning (iJET), 16(13), 29-39.

Heller, K. A. & Perleth, C. (2000). Kognitiver Fähigkeitstest für 4. bis 12. Klassen, Revision (KFT 4-12+R). Hogrefe, 2000.

Hmelo-Silver, C. E., & DeSimone, C. (2013). Problem-based learning: An instructional model of collaborative learning. In The international handbook of collaborative learning (pp. 370-385). Routledge.

Krell, M. (2015). Evaluating an instrument to measure mental load and mental effort using Item Response Theory.

Kuo, B.-C., Liao, C.-H., Pai, K.-C., Shih, S.-C., Li, C.-H., & Mok, M. M. C. (2019). Computer-based collaborative problem-solving assessment in Taiwan. Educational Psychology, 40(9), 1164-1185. https://doi.org/10.1080/01443410.2018.1549317

Limesurvey GmbH. (2012). LimeSurvey: An open source survey tool [Software]. Hamburg, Germany: LimeSurvey GmbH. Retrieved from http://www.limesurvey.org

Minkley, N., Kärner, T., Jojart, A., Nobbe, L., & Krell, M. (2018). Students' mental load, stress, and performance when working with symbolic or symbolic–textual molecular representations. Journal of Research in Science Teaching, 55(8), 1162-1187.

Muraki, E. (1992). A generalized partial credit model: Application of an EM algorithm. ETS Research Report Series, 1992(1), i-30.

OECD. (2017). PISA 2015 collaborative problem-solving framework. Retrieved from https://www.oecd.org/pisa/pisaproducts/Draft%20PISA%202015%20Collaborative%20Problem%20Solving%20Framework%20.pdf.

R Core Team (2022). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Rost, M. (2021). Modelle als Mittel der Erkenntnisgewinnung im Chemieunterricht der Sekundarstufe I: Entwicklung und quantitative Dimensionalitätsanalyse eines Testinstruments aus epistemologischer Perspektive. Logos Verlag Berlin.

Slotta, J. D., & Linn, M. C. (2000). The knowledge integration environment: Helping students use the Internet effectively. Innovations in science and mathematics education: Advanced designs for technologies of learning, 193-226.

Sun, C., Shute, V. J., Stewart, A. E. B., Beck-White, Q., Reinhardt, C. R., Zhou, G., Duran, N., & D'Mello, S. K. (2022). The relationship between collaborative problem solving behaviors and solution outcomes in a game-based learning environment. Computers in Human Behavior, 128. https://doi.org/10.1016/j.chb.2021.107120

Tang, P., Liu, H., & Wen, H. (2021). Factors Predicting Collaborative Problem Solving: Based on the Data From PISA 2015. Frontiers in Education, 6. https://doi.org/10.3389/feduc.2021.619450

Wang, L. (2016). Assessment and evaluation of key competences for student development-lessons from collaborative problem solving of PISA2015. Glob. Educ, 45, 24-30.