Learning Chemistry through Collaboration: A Wizard-of-Oz Study of Adaptive Collaboration Support[1]

Bruce M. McLaren[1], Nikol Rummel[2], Niels Pinkwart[3], Dimitra Tsovaltzi[1], Andreas Harrer[4], Oliver Scheuer[1]

[1]Deutsches Forschungszentrum Für Künstliche Intelligenz (DFKI), Germany
[2]Albert-Ludwigs-Universität Freiburg, Germany
[3]Technische Universität Clausthal, Germany
[4]Katholische Universität Eichstätt-Ingolstadt, Germany
bmclaren@dfki.de

Abstract. Chemistry students often learn to solve problems by applying well-practiced procedures, but such a mechanical approach is likely to hinder *conceptual* understanding. We have developed a system aimed at promoting conceptual learning in chemistry by having dyads collaborate on problems in a virtual laboratory (VLab), assisted by a *collaboration script*. We conducted a small study to compare an adaptive and a non-adaptive version of the system, with the adaptive version controlled by a human wizard. Analyses showed a tendency for the dyads in the adaptive condition to collaborate better and to have better conceptual understanding. We present our research framework, our collaborative software environment, and results from the wizard-of-oz study.

1. Introduction

How can we get chemistry students to solve problems conceptually rather than simply applying mathematical formulas? Students tend to struggle with transfer problems slightly different from those illustrated in a textbook, because they do not grasp the underlying concepts and, often times, prefer simply to apply algorithms [2]. On the other hand, research in chemistry education has suggested that *collaborative* activities can improve conceptual learning [3] and increase student performance and motivation [4]. While there have been few controlled experiments investigating the benefits of collaborative learning in chemistry, evidence that collaboration is beneficial exists in other disciplines, such as physics [5] and algebra [6]. This past work led us to investigate the advantages of collaborative activities in chemistry learning.

Collaborative partners typically need prompting and/or guidance to engage in productive interactions; thus, our approach is to support students with collaboration scripts, i.e., providing prompts and scaffolds that guide

---

[1] This paper is derived from a paper to be presented at the main ECTEL-08 conference [1].

students through their collaboration (e.g., [7]). Furthermore, students may be overwhelmed by the concurrent demands of collaborating, following script instructions, and trying to learn [8, 9], or, on the flipside, more advanced learners may not require as much support. We therefore hypothesize that adaptive collaboration support – i.e. scripting that changes over time based on characteristics of and actions taken by the learners – will increase the likelihood that students will attain conceptual chemistry knowledge. Some prior research has pointed toward the benefits of such adaptive support [10]. Our initial approach, discussed in this paper, is to provide adaptive collaboration support through a human wizard. Once we better understand how adaptive support benefits chemistry learners, we will automate the adaptive support.

2.   Technology Support for Chemistry Learning

Our approach entails student dyads collaborating on problems in a virtual chemistry laboratory. In particular, we use the *VLab*, a web-based software tool that emulates a chemistry laboratory [11]. We have extended the VLab software so that it is collaborative; that is, students on different computers can share and solve problems in the same VLab instance.
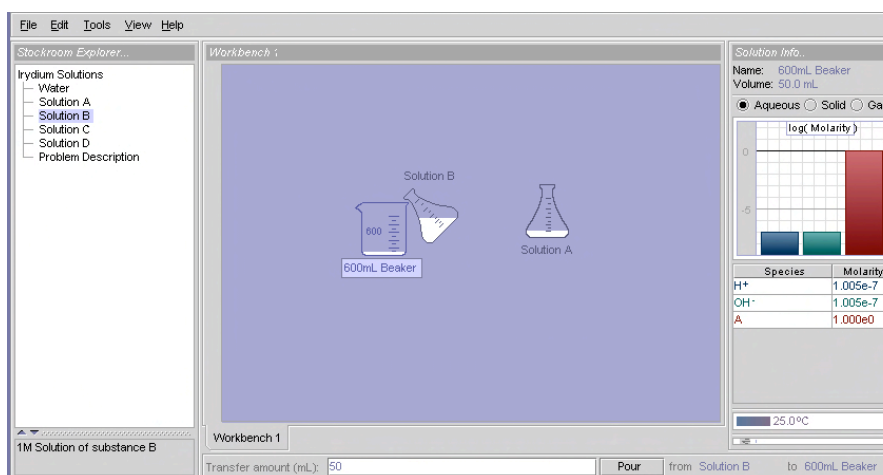

Fig. 1. Screenshot of the VLab

The VLab provides virtual versions of many of the physical items found in a real chemistry laboratory, including chemical solutions, beakers, bunsen burners, etc. and has meters and indicators for real-time feedback on substance characteristics, such as molarity. In Figure 1, two substances (Solution A and Solution B) have been dragged into the VLab workspace

(see the middle). 50 mL of Solution A has been poured into a separate 600mL beaker; 50 mL of Solution B is about to be mixed with this substance. The substance types and molarity within each container can be seen in the display on the right side of Figure 2 for a selected container. The idea behind the VLab is to provide students with an authentic laboratory environment in which they can run experiments, evaluate the changes that occur when mixing substances, very much like they would do in a real chemistry lab.
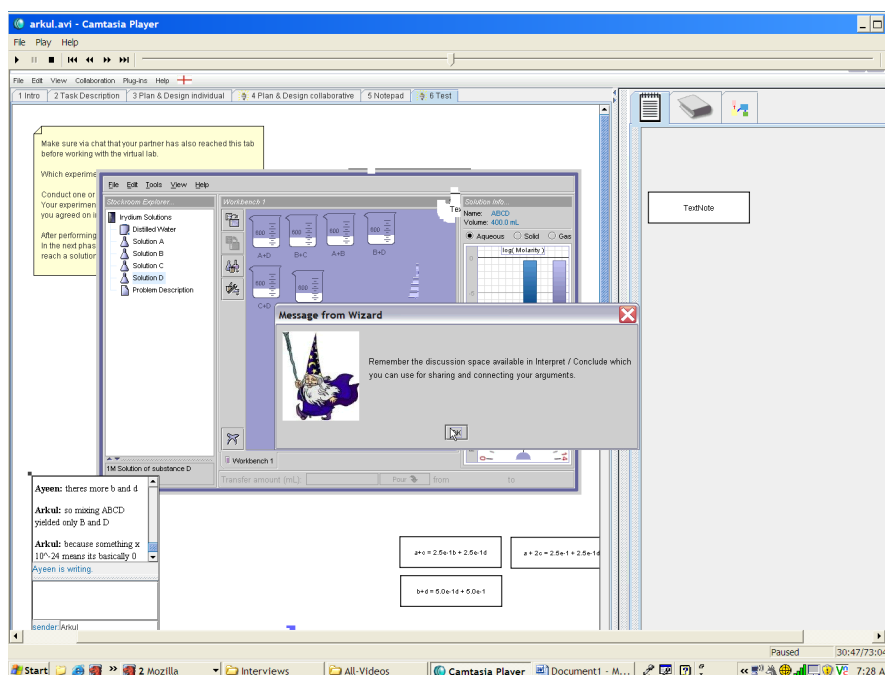

Fig. 2. A screenshot of the computer-based CoChemEx script, showing the *Test* tab

To support collaboration with the VLab, we integrated the software into an existing collaborative environment called FreeStyler [12], a collaborative software tool that is designed to support "conversations" and shared graphical modeling facilities between collaborative learners on different computers. Figure 2 shows the VLab in the middle, embedded in the Free-Styler environment. FreeStyler supports inquiry and collaboration scripts, using a third-party scripting engine, the CopperCore learning design engine. As explained in more detail in [12], the scripting engine can control the tools available within FreeStyler (e.g., chat, argumentation space, or VLab) for each phase of a learning activity. For the study described in this

paper, we complemented the FreeStyler scripting process with a human supervising the collaborating students and giving advice in a Wizard-of-Oz fashion2. The human wizard was able to send text messages and pictorial information directly to the collaborators (e.g., see the dialog in the middle of Figure 2).

3. Pedagogical Approach and Script

Our approach to scripting is to guide the collaborating students through phases of scientific experimentation and problem solving. More specifically, we base our script on the kinds of cognitive processes identified as typically used by experts when solving scientific problems experimentally, such as orientation, planning, and evaluation [cf 14]. Our experience with an initial version of the script, which prompted students to closely follow such a "scientific experimentation script," seemed to be too complex for students and thus led us to a simplification. The main steps of the current script, illustrated at the top of Figure 2 as tabs, are: *Plan & Design*, in which the dyads discuss their individual plans and agree on a common plan, *Test*, in which the collaborative experimentation in VLab takes place, and *Interpret & Conclude*, for discussing the results found in VLab and drawing conclusions. We also now guide students through the various steps in a less rigid manner to avoid overwhelming them with too much structure. The current approach gives general guidance on the script and provides prompts on solving VLab problems collaboratively. This approach is reminiscent of White *et al* [15] and Van Joolingen *et al* [16], which scaffold students as they collaboratively solve scientific problems. However, our focus is different: we are interested in how such an approach can be automated and if such support can bolster specifically the collaborators' *conceptual knowledge*.

In our approach, students are guided by static instructions in each tab. The first tab is the Task Description. The tabs Plan & Design Individual and Notepad allow each of the participants to record private notes and ideas using free-form text, in preparation for collaboration. The tabs Plan & Design Collaborative, Test, and Interpret & Conclude implement the script to guide the students' collaborative experimentation. Finally, in the tab Check Solution students submit their solutions and get error feedback. In the first cycle, the students are requested to follow this pre-specified order of steps

---

2 In a Wizard-of-Oz experiment, the participant interacts through an interface that includes a human "wizard" simulating possible system behavior [13]. The Wizard-of-Oz methodology is commonly used to investigate human-computer interaction in systems under development, with the goal of eventually automating the wizard's actions within the system.

and to click a "done" button to activate the next tab. After the first cycle, all tabs are available for a more open exploration.

Collaborating students work on separate computers and have access to a number of tools. The VLab (in the middle of Figure 2) is the basic experimental tool and the core collaborative component; it is situated in the Test tab. The chat window in the lower left of Figure 2 allows free-form communication between the students in the Test tab, as a way to explain, ask/give help, and co-construct conceptual knowledge. (Of course, as pointed out by one reviewer of this paper, providing the chat does not in and of itself lead to explanations or knowledge co-construction; such behavior must be supported and scaffolded through appropriate prompting, such as what the wizard provides in the current version of the system and automated support might later provide.) An argument space is available in the tabs Plan & Design collaborative and Interpret & Conclude. This allows the collaborators to discuss their hypotheses and results and to communicate general ideas, so as to promote students' conceptual understanding of the experimental process. It provides students with different shapes and arrows of different semantics for connecting the shapes. By using these components, students can make claims, provide supporting facts, and make counter-claims. In the shapes we provide sentence openers to prompt the argumentation, such as "I think that the main difference between our approaches to the problem is..." The argument space has the potential to allow students to reflect on each other's ideas [17]. Finally, a glossary of chemistry principles is available to the collaborating students at all times.

A human wizard provides adaptive support using a flowchart to observe and recognize situations that require a prompt, and to choose the appropriate prompt. The situations are defined by observable problematic behaviors in the tab where the activity currently takes place, either with regard to the collaboration (bad collaborative practice, e.g. ignoring requests for explanations), or with regard to following the script (bad script practice, e.g. moving to the next tab without coordinating with the partner). The wizard prompts are focused on providing collaboration support. We reviewed the literature on collaborative learning and developed a *top-down* version of the flowchart of prompts [5, 18] and then wrote collaboration prompts based on a *bottom-up* analysis of results from our earlier small-scale study. More specifically, we focused our adaptive feedback on prompting for communication (e.g., reminding to give and request explanations and justifications) and prompting after poor communication (e.g., reminding not to ignore requests for explanations or to contribute to the activities equally). This was a reaction to results from the small-scale study, in which it was revealed that students did not exhibit the right amount and kind of communication. A few prompts specific to our script remind students which tabs to use for their activities. Finally, domain-specific hints

are used as "dead-end prevention" in case students submitted a wrong solution. Two incorrect submissions are allowed; after that no more attempts are possible.

Figure 3 shows an example of one of the prompts in our flowchart, along with both the bottom-up ("Observed behavior") and top-down ("Theoretical foundation") branches of the flowchart that lead to this prompt. The entire flowchart, as well as discussion of its many details, is provided in [19].
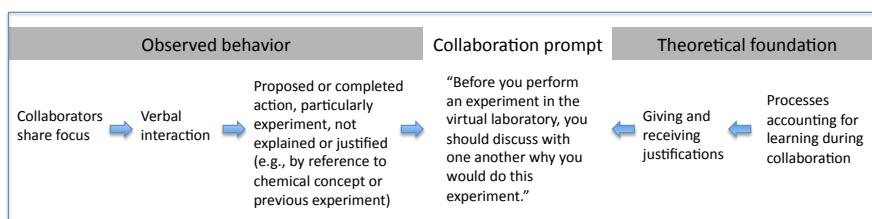

Fig. 3. Example of a Collaboration Prompt, arrived at by the wizard through observed behavior, but supported by theory

4. Wizard-of-Oz Study

We performed a small between-subjects wizard-of-oz study to test our computer-based collaborative learning environment and to refine the scripting approach based on an in-depth analysis of the data, with a focus on the adaptive aspects of the script. Our goal was to get a preliminary impression whether an adaptive system might lead to conceptual learning gains. Our study had 3 dyads per condition, with all subjects being university students. The experimental process followed a standard pre-test – intervention – post-test paradigm. In the intervention phase, two conditions were implemented: one using the standard version of the script, one using the adaptive version of the script. The adaptive social prompts by the human wizard were unique to the adaptive condition. Both conditions had to solve two chemistry problems of average difficulty. After the intervention phase a post-questionnaire and a post-test were administered. The post-test was equivalent to the pre-test, but included additional conceptual questions.

Quantitative Results. The results showed a tendency toward better conceptual understanding in the adaptive condition. Two conceptual questions were asked in the post-test for each of the problems. The concepts tested were all central to the tasks students encountered in the VLab. With a highest possible score of 6 points, the mean of the adaptive condition was M=4.6 (SD 1.63) whereas the non-adaptive condition scored in average

M=3.5 (SD 2.81). Due to the small sample size we did not perform further statistical analyses. An interesting result from the analysis of the post-questionnaire was that the adaptive condition reported a stronger impression that they did not have an equal chance to participate in solving the problems (on a 6-point Likert scale: Mad=5.16, SDad=1.36 vs. Mnon-ad=2, SDnon-ada=.6), although our process analysis revealed that such a difference is not real. On the other hand, this could be a cue that the wizard prompts to participate equally raised the participants' awareness of instances when participation was not equal. That is a desirable effect especially if it leads to corresponding attempts to balance participation.

Table 1. Summary of the process analysis of the script and collaboration practice.

Analysis Category	Number of Occurrences			
	Adaptive		Non-adaptive	
	M	SD	M	SD
Good script practice, e.g., coordinated actions in tab	6.33	2.51	5	2.64
Bad script practice, e.g., uncompleted actions	4.33	3.21	7.33	2.3
Good collaborative practice, e.g., ask for and give explanations	5.66	1.15	3	1
Bad collaborative practice, e.g., not explaining actions	2	1	1.66	1.15
Good reaction to a wizard message, e.g., improved practice after	8	4.58	(does not apply)	
Bad reaction to a wizard message, e.g., message has no apparent effect	6	4.7	(does not apply)	

Progress of individual dyads	Ad-Dyad-1: improved	Ad-Dyad-2: improved	Ad-Dyad-3: improved (slightly)	Non-Ad-Dyad-1: deteriorated	Non-Ad-Dyad-2: deteriorated (slightly)	Non-Ad-Dyad-3: stable

Process analysis of Study 2 Data The process analysis of the screen recordings of the collaborations revealed interesting differences between the two conditions, as shown in the summary in Table 1. Three members of our research team annotated different screen recordings independently. We counted the number of occurrences of good and bad script practice per dyad, that is, student's behavior relating to the script features (tab structure, argument space, and instructions). We also counted good and bad collaborative practice, that is, the kind of behavior expected and fostered by the prompts in the wizard's flowchart.

As shown in Table 1, there was a big difference between conditions and for both problem-solving sessions in the aggregated occurrences of "good script practice" and "good collaborative practice" in favor of the

adaptive dyads. "Bad script practice" was also considerably less frequent in the adaptive condition. However, the adaptive dyads showed slightly worse collaborative practice than the non-adaptive dyads. The category "Progress of individual dyads," at the bottom of Table 1, is a qualitative overall evaluation of each dyad as perceived by the annotators. It is a summary of the script and collaboration practice and the reaction to the wizard messages in the adaptive condition, per dyad. Notice that the adaptive dyads all improved, while the non-adaptive dyads remained stable or deteriorated. By "deteriorated" we mean that the non-adaptive dyads started out collaborating very well, but towards the end of the intervention period these dyads appeared to be discouraged and not seriously trying to solve the problems.

A detailed qualitative analysis of the deterioration of collaboration by the non-adaptive dyads, as well as analysis of other categories shown in Table 1, within the context of an actual dyad session, is provided in [1].

5.   Future Steps

Our initial results are only preliminary, based on a small sample of student dyads. Nevertheless, we see great promise in our approach. In the next steps, we will improve the script, making movements between tabs more flexible. The practical need to move between phases of experimentation, but our system's constraint against it (even though they were not strictly enforced), appeared to hinder the students on occasion. Also most of the ignored prompts were the ones that insisted on the use of the tabs in the prescribed sequence, another indication that this aspect should be changed.

We also plan to automate the feedback, which is currently provided by the human wizard based on specific student actions. The general idea is to use our flowchart as the "backbone" for development of the automated feedback approach, but adding techniques for automatically identifying situations, such as that illustrated in Figure 3. Of course, we will also carefully analyze which prompts in our flowchart appeared to lead to better (or worse) collaboration, or unwanted/unhelpful interruption to student progress, and update the flowchart accordingly. For the *Test* tab in particular, we plan to explore action analysis (e.g. [20]), extending Mühlenbrock's approach by analyzing VLab actions with machine learning techniques to identify situations in which prompts are necessary.

References

[1]   Tsovaltzi, D., Rummel, N., Pinkwart, N., Scheuer, O., Harrer, A., Braun, I., McLaren, B.M.: CoChemEx: Supporting Conceptual Chemistry Learning via Computer-Mediated Collaboration Scripts. In the Proceedings of the Third European Conference on Technology Enhanced Learning (ECTEL-08). (2008)

[2]   Gabel, D. L., Sherwood, R. D., Enochs, L.: Problem-Solving Skills of High School Chemistry Students. Journal of Research in Science Teaching 21 (2), 221--233 (1984)

[3]   Kozma, R. B.: The use of multiple representations and the social construction of understanding in chemistry. In M. Jacobson & R. Kozma (eds.), Innovations in science and mathematics education: Advanced designs for technologies of learning, pp. 11--46. Mahwah, NJ: Erlbaum. (2000)

[4]   Sumfleth, E., Rumann, S., Nicolai, N.: Kooperatives Arbeiten im Chemieunterricht [Cooperative Work in the Chemistry Classroom]. Essener Unikate 24, 74--85 (2004)

[5]   Hausmann, R. G., Chi, M. T. H., Roy, M.: Learning from Collaborative Problem Solving: An Analysis of Three Hypothesized Mechanisms. In: K. D. Forbus, D. Gentner , T. Regier (eds.) 26th annual Conference of the Cognitive Science Society, pp. 547--552. Mahwah, NJ, Lawrence Erlbaum. (2004)

[6]   Diziol, D., Rummel, N., Spada, H., McLaren, B. M.: Promoting Learning in Mathematics: Script Support for Collaborative Problem Solving with the Cognitive Tutor Algebra. In: C. A. Chinn, G. Erkens & S. Puntambekar (eds.), Mice, minds and society, CSCL 2007 8(I), 39--41. (2007)

[7]   Kollar, I., Fischer, F., Hesse, F. W.: Collaboration scripts - a conceptual analysis. Ed. Psych. Review, 18 (2), 159-185. (2006)

[8]   Dillenbourg, P.: Over-scripting CSCL: The Risks of Blending Collaborative Learning with Instructional Design. In: P. A. Kirschner (eds.) In: Three worlds of CSCL. Can we support CSCL, pp. 61—91. Heerlen: Open Universiteit Nederland (2002)

[9]   Rummel, N., Spada, H., Hauser, S.: Learning to Collaborate in a Computer-Mediated Setting: Observing a Model Beats Learning from Being Scripted. In: S. A. Barab, K. E. Hay, D. T. Hickey (eds.) Proceedings of the Seventh International Conference of the Learning Sciences, pp. 634--640. Mahwah, NJ: Lawrence Erlbaum Associates (2006)

[10]   Kumar, R., Rosé, C. P., Wang, Y. C., Joshi, M., Robinson, A.: Tutorial Dialogue as Adaptive Collaborative Learning Support. In: Proceedings of Artificial Intelligence in Education (2007)

[11] Yaron, D., Evans, K., Karabinos, M.: Scenes and Labs Supporting Online Chemistry. Paper presented at the 83rd Annual AERA National Conference, (2003)

[12] Harrer, A., Malzahn, N., Hoeksema K., Hoppe U.: Learning Design Engines as Remote control to learning support environments. Journal of Interactive Media in Education, Tattersall, C., Koper, R. (eds.) Advances in Learning Design, (2005)

[13] Bernsen, N. O., Dybkjær, H., Dybkjær, L.: Designing Interactive Speech Systems. From First Ideas to User Testing. Springer Verlag, (1998).

[14] De Jong, T., van Joolingen W.R. Scientific Discovery Learning with Computer Simulations of Conceptual Domains. Review of Ed. Research, 68(2), 179-201. (1998).

[15] White, B. Y., Shimoda, T. A., Frederiksen, J. R.: Enabling Students to Construct Theories of Collaborative Inquiry and Reflective Learning: Computer Support for Metacognitive Development. International Journal of Artificial Intelligence in Education. 10, 15--182 (1999)

[16] van Joolingen, W. R., de Jong, T., Lazonder, A. W., Savelsbergh, E., Manlove, S.: Co-lab: Research and Development of an On-Line Learning Environment for Collaborative Scientific Discovery Learning. Computers in Hum. Behavior, 21, 671--688 (2005)

[17] de Groot, R., Drachman, R., Hever, R., Schwarz, B., Hoppe, U., Harrer, A., De Laat, M., Wegerif, R., McLaren, B. M., Baurens, B.: Computer Supported Moderation of E-Discussions: the ARGUNAUT Approach. In: the Proceedings of Computer-Supported Collaborative Learning (CSCL-07), 165--167 (2007)

[18] Weinberger, A., Stegmann, K., Fischer, F., Mandl, H.: Scripting Argumentative Knowledge Construction in Computer-Supported Learning Environments. In: Fischer, F., Kollar, I., Mandl, H. & Haake, J.M. (eds.) Scripting Computer-Supported Collaborative Learning. Cognitive, Computational and Educational Perspectives. New York: Springer (2007)

[19] Braun, I.: Promoting Chemistry Learning Through Scripted Collaboration: Structural and Adaptive Support for Collaboration in a Computer-Supported Learning Environment. Diplomarbeit (Master's Thesis), Albert-Ludwigs-Universität Freiburg, Germany, June, 2008 (2008)

[20] Mühlenbrock, M.: Shared Workspaces: Analyzing User Activity and Group Interaction. In: Hoppe, H. U., Ikeda, M., Ogata, H., Hesse, F. (eds.) New Technologies for Collaborative Learning, Computer-Supported Collaborative Learning Series, Kluwer (2004)