Case Study Analysis on Collaborative Problem Solving Using a Tangible Interface

Dimitra Anastasiou(✉) and Eric Ras

IT for Innovative Services, Luxembourg Institute of Science and Technology, 5, avenue des Hauts Fourneaux, 4362 Esch-sur-Alzette, Luxembourg
{dimitra.anastasiou,eric.ras}@list.lu

Abstract. In our research within the "Gestures in Tangible User Interfaces" project we examine how gestural and speech behaviour while solving a collaborative problem on tangible interfaces has an effect on 21st Century skills. This paper presents an explorative case study where 15–22 year old pupils of a public school in Luxembourg solved a collaborative problem using a tangible user interface (TUI). The main goal of the study was to investigate the frequency and kinds of gestures and the link between the usage of gestures with collaboration and problem solving. At the end of the study, participants had to fill in three standard post-test questionnaires about workload and usability, the results of which are presented here and show that the use of a TUI as an educational medium was regarded as a straightforward and simple solution that many people could learn to use very quickly. This promising result is in line with our research vision, which is to include TUIs in a collaborative setting educational environment.

1 Introduction

Collaborative problem solving (ColPS) is a relevant competency which is extremely needed nowadays at workplace, but also in public and personal life. According to Tager (2013), the terms "collaborative problem solving", "cooperative work", and "group work" are used interchangeably in the education research literature to mean similar constructs. We follow the definition of Griffin et al. (2012), where ColPS refers to the abilities to recognize the points of view of other persons in a group; contribute knowledge, experience, and expertise in a constructive way; identify the need for contributions and how to manage them; recognize structure and procedure involved in resolving a problem; and as a member of the group, build and develop group knowledge and understanding.

In order to measure ColPS skills, our research project goal is to analyse the participants' gestural and speech behaviour while solving a ColPS scenario on a tabletop. Co-speech gestures are a very intuitive and expressive communication means and communication plays a crucial role in ColPS settings. Gestures have been often analysed in human-human conversational settings and also many gesture-based applications in human-computer interaction have been developed, but a systematic analysis of gestures on the application of TUIs is still lacking. Our aim is to incorporate gesture analysis as an assessment construct within the context of technology-based assessment (TBA).

D. Joosten-ten Brinke and M. Laanpere (Eds.): TEA 2016, CCIS 653, pp. 11–22, 2017.
DOI: 10.1007/978-3-319-57744-9_2

The objectives of the explorative case study are to evaluate whether the developed e-assessment scenario is appropriate for the target group, keeping the pupils, on one hand, engaged and motivated, but on the other hand, with a quite high level of cognitive mental and temporal load. The results of the pilot study led to necessary improvements of the scenario and the assessment approach before conducting the main evaluation study with more participants.

The next section elaborates on the existing frameworks for ColPS and complex problem solving (CPS). Section 3 describes the design of the pilot study and Sect. 4 presents the data collection as well as the discussion of the analysed data with regards to workload, usability, and user acceptance of the TUI in an educational environment.

2 Related Frameworks

Noteworthy are the definitions not only of ColPS, but also CPS competences by the international educational Programme for International Student Assessment (PISA). The PISA 2015 CPS framework[1] identified three main *collaborative competences*:

1. Establishing and maintaining shared understanding;
2. Taking appropriate action to solve the problem;
3. Establishing and maintaining team organization.

The framework included also the so-called *CPS competencies*:

1. Exploring and understanding;
2. Representing and formulating;
3. Planning and executing;
4. Monitoring and reflecting.

Currently, in practice, PISA instantiates ColPS by having problem solving units including chat-based tasks, where students interact with one or more computer agents or simulated team members, to solve the problem at hand.

In our opinion, e-assessment of ColPS should rather involve a face-to-face setting with collocated people (see discussion in Petrou and Dimitracopoulou 2003; Dillenbourg 1999). Dillenbourg (1999) takes the unsatisfactory for him definition of *collaborative learning* as "a situation in which two or more people learn something together" and interprets in a different way the elements included in this definition. For example, whether "two or more" means a small group or a few hundred or "together" means face-to-face or computer-mediated. Our contribution towards collaborative learning is to conduct case studies where groups of three pupils solve a complex problem collaboratively on a new educational ICT medium, which is a tangible tabletop. One long-term objective of our project is to provide Interaction Design guidelines for ColPS using a TUI either for high-stake assessments (e.g. future PISA programmes) or in other settings where the aspect of learning by means of formative assessment is in focus. The well-established collaborative and CPS competencies

[1] https://www.oecd.org/pisa/pisaproducts/Draft%20PISA%202015%20Collaborative%20Problem%20Solving%20Framework%20.pdf, 06.06.16.

defined by PISA are still in place in our studies, but the distinct difference is that we offer a *collaborative* setting for ColPS rather an agent-based chat interaction, as in PISA.

Both collaboration and ICT technology are key elements in the Partnership for 21st Century Learning[2]. According to this framework (see Fig. 1), 21st century skills are different than 20th century skills primarily due to the emergence of very sophisticated ICT technologies. For example, the types of work done by people – as opposed to the kinds of labour done by machines – are continually shifting as computers and telecommunications expand their capabilities to accomplish human tasks.

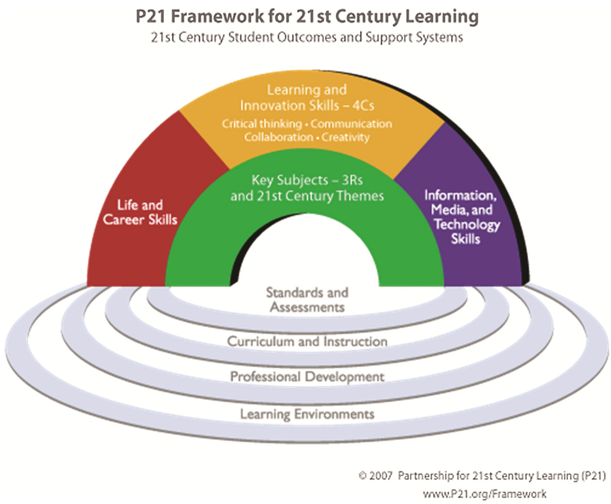

Fig. 1. 21st century student outcomes (arches of the rainbow) and support systems (pools at the bottom) according to the framework for 21st century learning

Binkley et al. (2012) compared a number of available assessment frameworks for 21st century skills which have been developed around the world and analysed them based on their learning outcomes in measureable form. Based on this analysis, they developed the so-called KSAVE model (*K*nowledge, *S*kills, and *A*ttitudes, *V*alues and *E*thics) with the following ten important skills (Table 1):

Csapó et al. (2012) reviewed the contribution of ICT technologies to the advancement of educational assessment. They described main tendencies for computer-based assessment in four continents (Asia, Australia, Europe, and USA) and with regards to challenge of assessment, they talked about five general situations, each of which poses different implications for the role that technology might play in assessment of twenty-first-century skills:

i. The first is characterized by domains in which practitioners interact with the new technology primarily using specialized tools, if they use technology tools at all;

[2] http://www.p21.org/, 06.06.16.

Table 1. The ten 21st century skills based on the KSAVE model

Ways of Thinking
1. Creativity and innovation
2. Critical thinking, problem solving, decision making
3. Learning to learn, metacognition
Ways of Working
4. Communication
5. Collaboration (teamwork)
Tools for Working
6. Information literacy (includes research on sources, evidence, biases, etc.)
7. ICT literacy
Living in the World
8. Citizenship – local and global
9. Life and career
10. Personal and social responsibility – including cultural awareness and competence

ii. The second situation is characterized by those domains in which, depending upon the preferences of the individual, technology may be used exclusively or not at all;
iii. The third situation is defined by those domains in which technology is so central that removing it would render it meaningless;
iv. The fourth situation relates to assessing whether someone is capable of achieving a higher level of performance with the appropriate use of general or domain-specific technology tools than would be possible without them;
v. The fifth situation relates to the use of technology to support collaboration and knowledge building.

In our user studies we focus on the fifth domain of technology according to Csapó et al. (2012) and the two skills of *Ways of Working* by Binkley et al. (2012) which are communication and collaboration.

3 Pilot User Study

In this section we refer to the general set-up of the study, particularly the demographic and social data of the participants (3.1), the micro-world scenario defining the task of the pupils (3.2), as well as the technical setup (3.3).

3.1 Participants

In March 2016 we ran a pilot case study with 15 participants in total split into 5 groups of three participants each. The students were between 15–22 years old (M = 17). Two groups were mixed gender (2 female & 1 male, 2 male & 1 female), two only female, and one only male. The language that participants talked to each other was Luxembourgish. The participants were coming sequentially and grouped as a team in the classroom where the study took place. They were advised to bring the consent forms which were given in advance, as both parents and/or legal representatives had to sign them too.

Apart from the consent form, the participants were not informed in advance in any way pertaining to the problem solving task they would deal with in this particular study.

3.2 Problem Solving Based on a Micro-World Scenario

In this subsection we refer both to the hardware and the software we used in order to design a collaborative problem solving scenario on a TUI. Our goal was that the selected scenario has a learning educational impact, but at the same time, is still enjoyable for the pupils of that age to keep them engaged in the scenario. The scenario does not require any previous knowledge from the pupils. This helps also to measure their learning effect independently after the user studies. Moreover, the overall task should not be difficult to solve, as it is not an IQ test or a typical exam *per se*; the goal of our project is to analyze their gestural and speech behavior while solving the ColPS on the tabletop.

As far as the hardware is concerned, the TUI employed for this pilot study is realised as a tangible tabletop (75 × 120 cm). Physical objects that have a specific reacTIVision[3] marker underneath can be recognized and manipulated on the table in order to explore different factors. reacTIVision is an open source, cross-platform computer vision framework for the tracking of fiducial markers attached onto physical objects, as well as for multi-touch finger tracking.

Our micro-world scenario was about building a whole power grid; this task is similar to tasks given in the PISA programme. The scenario is designed by means of a software framework TULIP (Tobias et al. 2015) which supports the instantiation of ColPS scenarios on TUIs via an XML file. In our pilot study, there was a picture from the Luxembourg city depicted on the TUI. The three pupils were provided with three physical objects (one object per participant) made out of cardboard, having a label A, B, and C on the top. These variables represented industrial facilities that produce electricity, i.e. a wind park, solar park, and a coal-fired power plant. The variables A, B, and C were used in order to prevent the potential previous knowledge of the pupils from solving the ColPS in a different way. The A, B, C objects could be placed, dragged, and rotated on the TUI. The factors that could be changed based on the placement and rotation of the objects[4] were: (i) power generation and (ii) CO_2 emission. The output values were data in gigawatt (gW) and in million tons (mT) of CO_2. In particular, when CO_2 increased, there was an animation on the TUI: the city got polluted and “foggier” due to the temperature rise created by the CO_2 emissions. Our micro-world scenario is based on several linear equations describing the relationship between the parameters. The scenario is shown in Fig. 2.

The instructions were orally given by an experimenter prior to the commencement of the study. After the instructions, the pupils had a practice session of three minutes to find out themselves how the objects affect the power generation and CO_2. After the practice session, they were asked to solve the following seven questions as presented in Table 2. The questions were visualized on the TUI itself and the experimenter manipulated a physical object to move from one question to the next one.

---

[3] http://reactivision.sourceforge.net/, 05.06.16.

[4] The dragging of the objects along the TUI did not have any effect.

Fig. 2. Collaborative problem solving on a tabletop

Table 2. Questions during ColPS in our user study

Q1	Which facility can produce the maximum power?
Q2	Which facility produces pollution (CO_2)?
Q3	Place one object on the table that produces the least power and no pollution.
Q4	Place as many objects as necessary on the TUI, so that power of 5,55 Gigawatt (GW) is produced in total.
Q5	Place as many objects as necessary on the TUI, so that power of 1,02 GW is produced in total.
Q6	Place as many objects as necessary on the TUI, so that pollution of 4 million tons (mT) pollution (CO_2) is produced in total.
Q7	Place as many objects as necessary on the TUI, so that power of 2,55 Gigawatt (GW) and maximum 2 mT CO_2 are produced in total.

3.3 Technical Set-Up

The required technical equipment was transported to the public school and two/three experimenters needed at least two hours to set it all up and calibrate the system. There were three cameras recording the study (see Fig. 3): (i) a Kinect 2.0[5] depth sense camera recording 3D data of the participants (ii) a standard video camera, (iii) a GoPro camera mounted on the ceiling. By means of Kinect 2.0 we developed an application that can automatically analyze object manipulation in real time with regards to *which object* has been manipulated *when*, *by whom* using *which hand*. A more detailed description of this application is outside the scope of this paper.

[5] http://www.xbox.com/en-US/xbox-one/accessories/kinect-for-xbox-one, 04.06.16.

Fig. 3. Technical set-up of the pilot user study in the classroom

4 Data Collection and Analysis

A multimodal corpus of video volume of about 2.5 h was collected for all five groups. The data collected through our user studies are of three types: (i) audio-visual recordings, (ii) 3D and logging data, and (iii) questionnaires. The audio-visual recordings are being annotated with the software ELAN (Wittenburg et al. 2006), a professional tool for the creation of complex annotations on video and audio resources. The Kinect 2.0 depth sense camera is used for recognition of the spatial position of the participants and their gestural interaction with the objects. Our software framework TULIP (Tobias et al. 2015) uses an abstraction layer to receive information from the computer vision framework and a widget-model based on model-control representations. In this paper we focus particularly on the third type of the collected data, the questionnaires. A general description of the questionnaires is presented in Subsect. 4.1 followed by the results of each questionnaire.

4.1 Questionnaires

After the end of the study, we handed three standard post-test questionnaires to the pupils, who filled them in at the classroom where the study took place. These three questionnaires were:

1. Nasa task load index (Hart 2006);
2. Attrakdiff (Hassenzahl 2004);
3. System Usability Scale (Brooke 1996).

The Nasa task load index is a subjective workload assessment tool with ratings on mental, physical, temporal demands, as well as own performance, effort and frustration. More precisely the questions included in the questionnaire are (Table 3):

Table 3. Nasa task load index questionnaire

Mental demand	How much mental and perceptual activity was required (e.g. thinking, deciding, calculating, remembering, looking, searching, etc.)?
Physical demand	How much physical activity was required (e.g. pushing, pulling, turning, controlling, activating, etc.)?
Temporal demand	How much time pressure did you feel due to the rate or pace at which the tasks or task elements occurred?
Performance	How successful do you think you were in accomplishing the goals of the task set by the experimenter (or yourself)?
Effort	How hard did you have to work (mentally and physically) to accomplish your level of performance?
Frustration level	How insecure, discouraged, irritated, stressed and annoyed versus secure, gratified, content, relaxed and complacent did you feel during the task?

The second questionnaire we used had the purpose to evaluate the usability and design of the micro-world scenario on the TUI. *Attrakdiff* is used to measure the attractiveness of a product or service in the format of semantic differentials. It consists of 28 seven-step items whose poles are opposite adjectives (e.g. "confusing - clear", "unusual - ordinary", "Good - bad"); see all items in Fig. 5. Each set of adjective items is ordered into a scale of intensity. Each of the middle values of an item group creates a scale value for pragmatic quality (PQ), hedonic quality (HQ) and attractiveness (ATT).

The third questionnaire is the System Usability Scale (SUS) questionnaire, a tool for measuring the usability of a product or a service. In our case "the product" was the tabletop employed in an educational environment. SUS consists of 10 items (see Table 4) with five response options (from strongly agree to strongly disagree).

Table 4. SUS questionnaire

Q1	I think that I would like to use this system frequently
Q2	I found the system unnecessarily complex
Q3	I thought the system was easy to use
Q4	I think that I would need the support of a technical person to be able to use this system
Q5	I found the various functions in this system were well integrated
Q6	I thought there was too much inconsistency in this system
Q7	I would imagine that most people would learn to use this system very quickly
Q8	I found the system very cumbersome to use
Q9	I felt very confident using the system
Q10	I needed to learn a lot of things before I could get going with this system

Results on Workload

The results (see Fig. 4) show that the workload on the performance was higher rated compared to the other categories. The performance was about "how successful do you think you were in accomplishing the goals". This result can be attributed to the fact that there was textual feedback ("You have reached the goal") about the task performance implemented in the ColPS scenario (see Table 2). The mental and temporal demand followed at the second and third place respectively. These results in combination with the results that the frustration level and the effort required were quite low show that the scenario selected for this age group was appropriate, keeping the pupils, on one hand, engaged and motivated, but on the other hand, with a certain level of cognitive mental and temporal load.

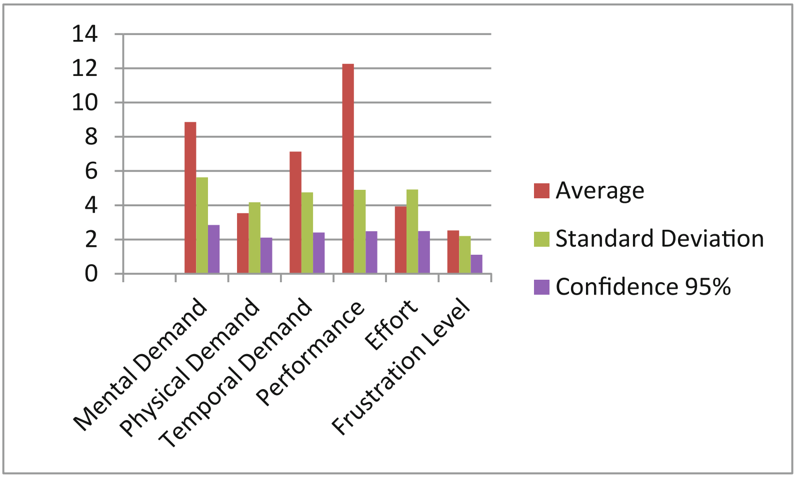

Fig. 4. Results of *the Nasa task load* questionnaire

Results on Attractiveness

The results of the Attrakdiff questionnaire are presented in Fig. 5. On the first picture (left), the various word pairs are presented along with their effect on pragmatic quality (PQ), hedonic quality (HQ) (including HQ-I and HQ-S) and attractiveness (ATT). The overall concept, solving a collaborative problem on the TUI, was ranked very *practical, creative,* and *good* by the participants. Striking is the result that although it was considered very *inventive*, it was also regarded as a *simple* and *straightforward* solution. This is a very positive result, showing that ICT technologies, and in our case the TUI, should not be intimidating, but rather employed to design straightforward scenarios where pupils can collaboratively solve a problem, get motivated, and even being provided with feedback. The portfolio presentation on the right side shows that the scenario balances between self-oriented and desired solution. This is the result of the scores of PQ (0.84) and HQ (1.30). The confidence rectangle shows that according to user consensus, the hedonic quality is greater than the pragmatic quality.

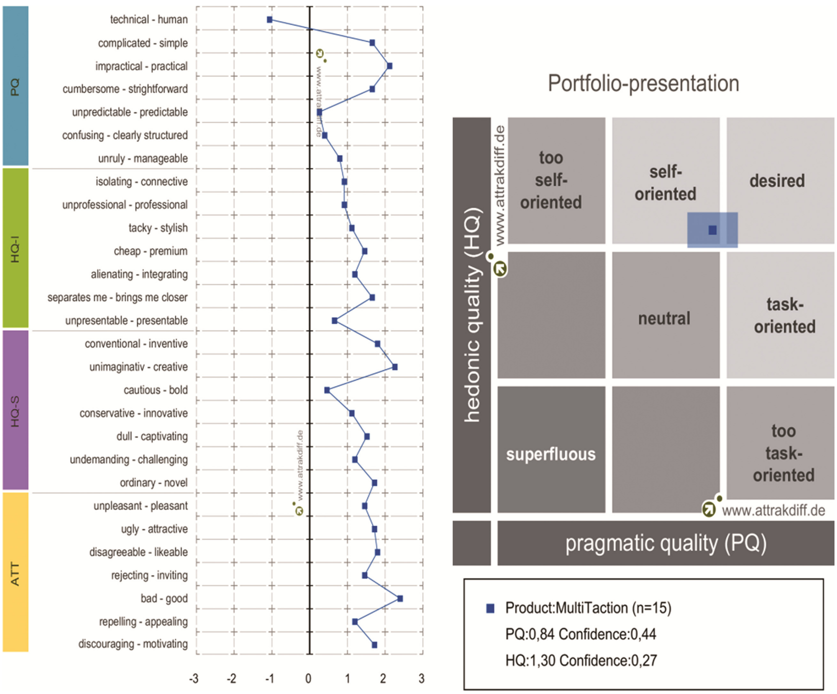

Fig. 5. Results of the *Attrakdiff* questionnaire

Results on Usability (SUS)

The resulting overall SUS score is 71.7, which is a very promising score. Figure 6 shows that the lowest SUS score was 45.0 given by the Participant 6 while the highest was 97.5

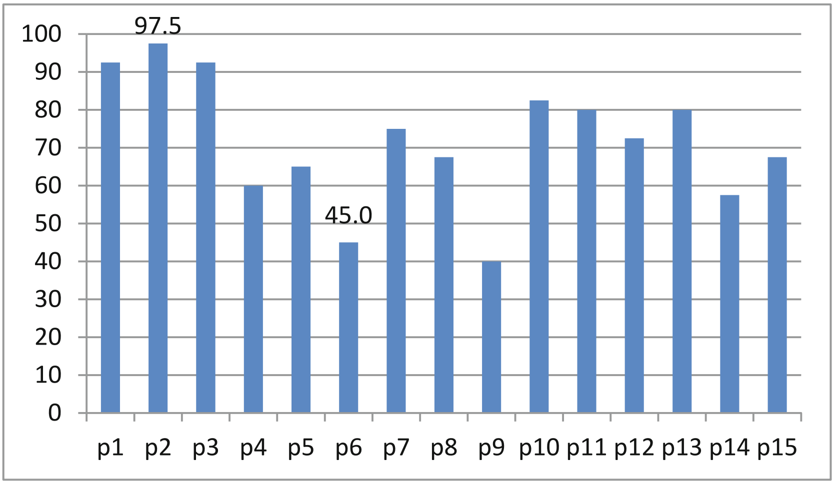

Fig. 6. Results of the *SUS* questionnaire

by Participant 2. Based on average scores, the question with the highest score was question 7 "I would imagine that most people would learn to use this system very quickly" which is a very optimistic and promising result particularly with regards to the applicability of TUIs in e-assessment in the future.

5 Conclusion and Future Prospects

An explorative case study with 15 participants was conducted in March 2016 in terms of the Marie Curie project GETUI. We spent about 2–3 months to design and implement an appropriate micro-world scenario on the TUI, which was employed in the study as a tangible tabletop, where physical objects can be used by the users to manipulate parameters in the scenario. In the design of the scenario, we took into account the PISA framework and its defined collaborative and complex problem solving competencies. Based on these requirements, our scenario was supposed to encourage and motivate users as well as to facilitate collaboration. These requirements were confirmed by the participants based on the results of the post-test questionnaires. Another significant result was that the micro-world scenario on the TUI was regarded as a simple and straightforward solution and in addition, that many people can learn to use the system very quickly.

As far as future prospects are concerned, we are currently annotating the video recordings. Our draft annotation scheme can be found at Anastasiou and Bergmann (2016). We distinguish between (i) physical or free hand movements (mainly pointing gestures) and (ii) manipulative gestures (dragging, rotating objects on the TUI). The former concern mainly human-human interaction, while the latter human-computer interaction. Speech transcription, conversational analysis as well as speech-gesture temporal alignment is planned at a later stage of the project. Moreover, a combination of our logging data from the TUI and the 3D data of the participants will help train our developed gesture recognition application which will decrease the time-consuming task of manual annotation in the long run.

Acknowledgment. This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no. 654477. We thank the director as well as the pupils of the school Lycée Technique d'Esch-Sur-Alzette for the collaboration.

References

Anastasiou, D., Bergmann, K.: A gesture-speech corpus on a tangible interface. In: Proceedings of Multimodal Corpora Workshop, International Conference on Language Resources and Evaluation (2016)

Brooke, J.: SUS – a quick and dirty usability scale. In: Jordan, P.W., Thomas, B., Weerdmeester, B.A., McClelland, I.L. (eds.) Usability Evaluation in Industry. Taylor & Francis, London (1996)

Binkley, M., et al.: Defining twenty-first century skills. In: Griffin, P., McGaw, B., Care, E. (eds.) Assessment and Teaching of 21st Century Skills, pp. 17–66. Springer, Dordrecht (2012)

Csapó, B., et al.: Technological issues for computer-based assessment. In: Griffin, P., Care, E. (eds.) Assessment and Teaching of 21st Century Skills, pp. 143–230. Springer, Dordrecht (2012)

Dillenbourg, P.: What do you mean by collaborative learning. In: Collaborative-learning: Cognitive and Computational Approaches, vol. 1, pp. 1–15 (1999)

Griffin, P., Care, E., McGaw, B.: The changing role of education and schools. In: Griffin, P., McGaw, B., Care, E. (eds.) Assessment and Teaching 21st Century Skills, pp. 1–15. Springer, Heidelberg (2012)

Hassenzahl, M.: The interplay of beauty, goodness, and usability in interactive products. Hum.-Comput. Inter. **19**(4), 319–349 (2004)

Hart, S.G.: NASA-task load index (NASA-TLX); 20 years later. In: Proceedings of the Human Factors and Ergonomics Society 50th Annual Meeting, HFES, pp. 904–908 (2006)

Petrou, A., Dimitracopoulou, A.: Is synchronous computer mediated collaborative problem-solving 'justified' only when by distance? Teachers' points of views and interventions with co-located groups, during everyday class activities. In: Wasson, B., Ludvigsen, S., Hoppe, U. (eds.) Designing for Change in Networked Learning Environments, vol. 2, pp. 441–450. Springer, Dordrecht (2003)

Tager, M.: Computer-based assessment of collaborative problem-solving skills: human-to-agent versus human-to-human approach, Research report (2013)

Tobias, E., Maquil, V., Latour, T.: TULIP: a widget-based software framework for tangible tabletop interfaces. In: Proceedings of the 7th ACM SIGCHI Symposium on Engineering Interactive Computing Systems, pp. 216–221 (2015)

Wittenburg, P., et al.: ELAN: a professional framework for multimodality research. In: Proceedings of the 5th Conference on Language Resources and Evaluation, pp. 1556–1559 (2006)

http://www.springer.com/978-3-319-57743-2

Technology Enhanced Assessment
19th International Conference, TEA 2016, Tallinn, Estonia, October 5-6, 2016, Revised Selected Papers
Joosten-ten Brinke, D.; Laanpere, M. (Eds.)
2017, X, 197 p. 48 illus., Softcover
ISBN: 978-3-319-57743-2