EAI Endorsed Transactions
on Serious Games

Research Article ICST.ORG

A case study on the design and development of mini-games for research methods and statistics

P. Van Rosmalen[1,*], E.A. Boyle[2], J. Van der Baaren[1], A.I. Kärki[3], Ángel del Blanco Aguado[4]

[1]Welten Institute, Faculty of Psychology and Educational Sciences, Open University of the Netherlands, P.O.Box 2960, 6401 DL Heerlen, The Netherlands
[2]School of Social Sciences, University of the West of Scotland, Paisley, United Kingdom
[3]Satakunta University of Applied Sciences, Maamiehenkatu 10, 28500 Pori, Finland
[4]Universidad Complutense de Madrid, Madrid, Spain

Abstract

Research methodology involves logical reasoning and critical thinking skills which are core competences in developing a more sophisticated understanding of the world. Acquiring expertise in research methods and statistics is not easy and poses a significant challenge for many students. The subject material is challenging because it is highly abstract and complex and requires the coordination of different but inter-related knowledge and skills that are all necessary to develop a coherent and usable skills base in this area. Additionally, while many students embrace research methods enthusiastically, others find the area dry, abstract and boring. In this paper we discuss the design and the first evaluation of a set of mini-games to practice research methods. Games are considered to be engaging and allow students to test out scenarios which provide concrete examples in a way that they typically only do once they are out in the field. The design of a game is a complex task. First, we describe how we used cognitive task analysis to identify the knowledge and competences required to develop a comprehensive and usable understanding of research methods. Next, we describe the games designed and how 4C-ID, an instructional design model, was used to underpin the games with a sound instructional design basis. Finally, the evaluation approach is discussed and how the findings of the first evaluation phase were used to improve the games.

Keywords: Research methods and statistics, Serious Games, Game Design, Cognitive Task Analysis, Instructional Design, 4C-ID.

Received on 14 January 2014, accepted on 02 June 2014, published on 26 August 2014

doi: 10.4108/ sg.1.3.e5

1. Introduction

The inductive and hypothetico-deductive reasoning that underlie research methodology involves logical reasoning and critical thinking skills that are core competences in developing a more sophisticated understanding of the world. These higher level thinking skills are required to tackle the ill-defined problems that we face in the 21st century. The EU 2020 Strategy "New Skills for New Jobs" (2010) emphasises the need to help students acquire skills for the kinds of jobs which will be available in the year 2020. It seems likely that the ability to understand and present convincing arguments and evaluate the quality of evidence are skills which will gain increasing importance in future.

However acquiring expertise in research methods and statistics is not easy and poses a significant challenge for many students (Tishkovskaya & Lancaster, 2010). The subject material is challenging because it is highly abstract and complex and requires the coordination of different but inter-related knowledge and skills that are all necessary to develop a coherent and usable skills base in this area. Students have to develop an understanding of how to formulate hypotheses, identify, define and operationalise relevant variables, select an appropriate design to examine links between variables, identify an appropriate sample of participants, identify relevant

[*]Corresponding author. Email:peter.vanrosmalen@ou.nl

ethical issues, select informative and suitable methods of data analysis, collect and analyse data, and interpret and discuss the findings. It is also challenging for teachers to generate interest in the subject. While some students embrace research methods enthusiastically, others find the area dry, abstract or complex. Teachers are continually looking for new ways of making research methods and statistics more appealing to their students.

There is growing evidence in the literature that serious games can be an effective tool to support learning (Joint Information Systems Committee JISC, 2007). Nadolski et al (2008) suggest that the use of serious games can be a useful tool for Higher Education Institutions to develop and deploy, to enhance the student experience and to assist them in achieving the intended learning outcomes. One advantage of games is that they can bridge the gap between theory and practice by allowing students to test out scenarios which provide concrete examples in a way that they typically only do once they are out in the field.

The CHERMUG project (www.chermug.eu) was conceived against this background of finding more effective and engaging ways to teach research methods and statistics and the recognition that a games-based approach might be useful in this respect. It is clear that games are highly motivating which is considered a main determinant of effective learning (Keller, 1983). The engagement that games provide could be especially important in generating interest in this notoriously abstract and difficult subject area. More importantly however, games offer activities that are highly consistent with modern theories of learning which emphasise that learning will be more effective when learners are actively engaged in carrying out activities that require both knowledge and skills and that reflect the kinds of real-world problems that are typical in that subject domain (Boyle, Connolly & Hainey, 2011). Learning will also be assisted when learners tackle a variety of problems from different perspectives that require them to use their knowledge in slightly different ways and where learners have access to support from more able individuals. Games have the potential to provide such activities, which will help the learner develop a usable knowledge base that they can deploy to solve real-world problems.

Despite the phenomenal growth of interest in serious games, there is still little systematic guidance concerning which kind of game is better for which purpose and so developing a game can be quite an experimental process. In developing a game for supporting complex problem solving, careful consideration needs to be given to the subject discipline, the content area, the player and the needs of the players in learning about the content area, pedagogy, the affordances of games and matching games to desired learning outcomes.

This paper presents a case study of the CHERMUG project which aimed to design a game for teaching research methods and statistics to nursing and social science students in a systematic way. The CHERMUG game design tasks included a literature review (Boyle, Manea, & Karki, 2013), a stakeholder and user requirements analysis (Boyle and MacArthur, 2013) and a cognitive task analysis (Boyle et al, 2012). This paper will describe different aspects of the game design and development in detail in the hope that it will be useful in helping others to think about issues that need to be tackled in developing games to support complex problem solving. We will focus in particular on the CTA carried out and explain how the results of the initial activities (literature review, user requirements analysis and CTA) were then used in the design and implementation of the game.

The literature review (Boyle, Manea, & Karki, 2013; Boyle et al, submitted), carried out around games, animations and simulations, could identify only a handful of papers reporting any empirical evaluation of games to teach research methods and statistics. Even fewer attempted to teach a full research cycle, with most concentrating on a specific stage. Operation ARA (Halpern et al, 2012), Martian Boneyards (Asbell-Clarke et al, 2012) and Ramler and Chapman's (2011) use of Guitar Hero are games where players had to propose a hypothesis, look for evidence and evaluate whether the evidence supported the hypothesis. The relative lack of games probably reflects the recent interest in the use of games for learning and the complexity and interdependence of knowledge in this area which may be difficult to gamify.

The user requirements analysis (Boyle and MacArthur, 2013) confirmed that, although at present there was little use of serious games, both higher education nursing students and staff in the collaborating countries were open towards the use of digital games as a component of a blended learning approach to teaching methods and statistics.

2. Cognitive Task Analysis (CTA)

A major constraint in introducing games into the curriculum is identifying the relevance of the game to the curriculum (Kirriemuir and McFarlane, 2004). A first step, therefore, in developing a game to support students in learning in specific curricular areas is to identify the skills and competences required. A technique which has been developed to help analyse the higher level cognitive functioning required in tackling complex tasks is Cognitive Task Analysis (CTA). CTA is defined as "the extension of traditional task analysis techniques to yield information about the knowledge, thought processes and goal structures that underlie observable task performance" (Chipman, Schraagen and Shalin, 2000, p. 3). CTA is typically carried out when knowledge about how a task is performed is uncertain.

On one level the knowledge which is required in developing an understanding of research methods and statistics is quite well known and is presented in many textbooks on the subject. What is not so certain however is the best way in which to present this knowledge to students. Lovett (1998) argued that CTA can help in describing the curriculum to be taught and decomposing the curriculum into the knowledge and sub-skills that students must learn. Lovett applied CTA to exploratory data analysis in statistics.

In CHERMUG, CTA was used to identify the cognitive skills, knowledge and competences required by students, to look at the way in which these are typically taught and the problems which students mostly experience in acquiring a comprehensive and usable understanding of research methodology and statistics.

In a period of 6 weeks 13 interviews were carried out (Boyle et al, 2012; Boyle, Van Rosmalen and Manea, 2013). The interviewees were selected based on their knowledge of and involvement with teaching research methods and statistics. Faculties covered included Nursing and Medicine, Social Science, Psychology, Learning Sciences plus one expert who advised and supported several faculties. The experts were spread over universities and professional universities, some of them covering both. The experts consulted were located in British (7), Dutch (3) and Romanian (3) higher education institutions. The participants received a briefing sheet in advance of their interview, providing an outline of the aims of CHERMUG project and the objectives of the CTA. Given the variety in participants' backgrounds, the briefing sheet also included a description of the different stages in the research methods cycle (research question, data collection, data analysis and discussion & conclusion) and a short description of three papers (Asbell-Clarke et al, 2012; Hummel et al, 2011; Hulshof, Eysink & de Jong, 2006) about games which had been identified as relevant either to the content area of the game or to the possible design of the game. The participants were asked in the semi-structured interview about their views on teaching research methods and about the research cycle as a reference point.

Overall, research methods are seen as a complex and challenging topic for students. A fundamental difference in perception between universities and professional universities concerns what level of competence or skills is expected with regard to research methods. The position taken by professional universities varied between 'being able to understand research methods' or 'being able to assess research papers in function of evidence based practice' to 'defining and executing a research plan'; it was expected that parts of the cycle, in particular the use of statistics, would be actively supported by a supervisor. Universities, on the other hand, in principle do expect that students can define and execute all steps of a research plan. Importantly for the design of the game, the research cycle presented was generally accepted by staff as providing a useful framework for presenting research methods to students. The most important suggestion shared by the respondents was that the teaching method should provoke research interest. Maybe the most important aim in teaching research methods is trying to increase student motivation and get them excited about research and interested in being part of the research community. Other suggestions mentioned are:

- **Challenge students** in their question definition and research design & show the dependencies of the full research cycle.
- **Experience the difference in research methods: qualitative or quantitative**. Several respondents suggested that it would be useful to illustrate differences between and advantages and disadvantages of qualitative and quantitative approaches.
- **Experience by being part of the experiment**. While an abstract dataset might be easier to apply across different domains, a 'personalised' dataset might be more appealing and intuitive.
- **Visualise**. Several respondents mentioned the potential of a game-based approach in helping students to visualise data: "To include in the game perspectives to enable 'understanding before analysing' data e. g. looking at extremes, making use of graphical representations as part of the process to get a global understanding of the data before making use of evaluative statistics."

Finally, a general conclusion was that, even more than expected, our target audience is extremely heterogeneous. There are different demands with regards to research methods depending of university type, country and domain. Nevertheless, beyond being clear about prerequisites, respondents did not appear to view the level of expertise at which the game was targeted as a problem. While initially it was thought that the game should be targeted at beginners, it seems that even a beginner's level could potentially be useful for students at all stages. The complexity of research methods and statistics suggests that even experts have areas where they might find it useful to revise their understanding.

2. The CHERMUG mini-games

2.1. Global Design

The CTA together with the results of the literature review and the requirements analysis resulted in a set of initial directives. The games should be targeted at beginners and be most useful for students who are taking an introductory module on research methods (or alternatively as a refreshment for more advanced students). The learner should only need a basic understanding of terminology and concepts used before the games can be played. The games should require no IT skills on the part of the tutor or the students, beyond being able to operate a web browser. Finally, the games should in particular raise interest in research methods and focus on two topics i.e. the research methods cycle and the distinction between qualitative and quantitative approaches to research.

Designing a serious game, however, is a complex operation and despite the growth of interest in serious games, there is still little systematic guidance concerning which kind of game is better for which purpose and how to assure a game fits the instruction required. The complexity of the field is clearly illustrated by, for instance, Connolly, Boyle, MacArthur, Hainey, & Boyle (2012) who in a recent review study classify games on genre, subject discipline and intended outcome. As a result, for developers the design and development of a game and for instructors the selection and application of a game can be quite an experimental process.

Only recently, there exist a number of frameworks which attempt to integrate the knowledge and experience with regard to education, games and software (see e.g. Harteveld, 2011; De Freitas et al, 2010; Pernin et al, 2012). These frameworks are important tools to assist in the design of serious games. However, they are as yet not fully matured nor investigated as, indirectly, shown in the limited evidence on the effectiveness of serious games and the apparent difficulty in assessing the educational merits of serious games (Connolly et al, 2012). Moreover, these frameworks do not necessarily fit with the background of teachers.

Table 1. The CHERMUG mini-games divided by their task classes and topics

	Task class (Topic) / Level		
	Qualitative	Quantitative (Chi-square)	Quantitative (T-test)
Level 1	Main differences between qual. & quant. Analysis	Gender & reward	Nationality & Mediterranean food
Level 2	Simulating a quantitative research study	Exercise program & drop-out	Gender & protein consumption
Level 3	Writing to a journal	Media consumption & obesity	Type of diet & weight loss
Level 4		Skipping meals & obesity	
Level 5		Nationality & body image	

Generally teachers have insufficient knowledge about games and their beneficial usage in classrooms (NFER, 2009). Educational games are considered fundamentally different from prevalent instructional paradigms (FAS, 2006). Williamson (2009) reports an urgent need for the training of teachers both at the initial training stage and the stages of continuous professional development, to pursue a better understanding of how to use games in their class-rooms as well as understanding the implications of games as cultural forms of young people's lives. The general impression is that games require complex technologies and that games are difficult to organise and to embed in a curriculum (Klopfler, Osterweil, & Salen (2009). The latter is of importance since the use of ICT, and games in particular, only tends to be successful if it closely fits with the existing teaching practice (Vier in Balans Monitor 2012, 2012).

A way to support the game design and to support the application of a game would be to build upon a proven framework which integrates a sound instructional foundation, fits with teachers' experiences and fits sufficiently with existing game principles. Huang and Johnson (2009) propose using the 4C-ID model. The underlying assumption of the 4C-ID model (Four Component Instructional Design) is that complex learning can be designed with the help of four interrelated components (Van Merriënboer, & Kirschner, 2012):

1. **Learning tasks**. Authentic, whole tasks preferably based on real-life tasks and organised in task classes with variation and increasing complexity.
2. **Supportive information**. Information that is supportive to the non-recurrent aspects of the tasks and explains how a domain is organised. This information is always available.
3. **Procedural information**. Information that is prerequisite to the recurrent aspects of tasks and instructs how to perform the routine aspects of a task. This information is available just-in-time and typically, stepwise will fade out when exercising with new tasks.
4. **Part-task practice**. Additional practice for routine aspects of learning tasks that require a high level of automation.

Together, the overall design focus is on the integration and coordination of different levels of learning tasks and as such fits very well with existing game design practice. Recent studies (Lukosch, Van Bussel & Meijer, 2012; Enfield, 2012) confirmed the applicability of the model for game design and their embedding in education. Giving the findings discussed above, the 4C-ID model was used to shape the global design of the CHERMUG games applying 4C-ID in the following way:

- **Authentic tasks**. A set of mini-games (table 1) was designed each based on an authentic and complete task dealing with a research problem starting from a global introduction and hypothesis to the discussion of the findings. Mini-games were chosen because they should fit easily into the curriculum and they should allow to quickly go through the main challenges of research thus helping to increase student' motivation and get them excited about research. The literature review suggested that teaching research methods and statistics is more successful when it is taught with content and examples which are relevant to the student and are grounded in real-life examples. The broad area of obesity was selected as a topic of general interest, in particular, to nurses and social scientist. There are many variables which are relevant to and impact on obesity and informal piloting in a class and with friends suggested that most lay people could quickly generate several factors which are related to obesity.
- **Task classes, variation and increasing complexity in task classes**. CTA and requirements analysis

indicated the importance of having access to practice with quantitative and qualitative research. For quantitative research two task classes were designed of similar intrinsic complexity, the first one for research problems to be addressed with a chi-square and the second with a t-test. To ease the use of the games all games follow a similar pattern, the games incrementally becoming slightly more complex (see table 1 levels) within their task class. The qualitative games are covered with only one task class with three games dealing with typical qualitative research scenarios. The first game introduces the student to qualitative research by studying the main differences between qualitative and quantitative research. In the next two games the students work through a simulated qualitative study. The three games become stepwise more complex starting with immediate feedback after each action on level 1 to only game completion feedback at level 3.

- **Supportive and procedural information**. 4C-ID proposes to progressively decrease the guidance for consecutive tasks in the same task class. However, since most of the games only take 10-15 minutes supportive information is expected to be offered in advance. Moreover, any procedural information required is implicit through the rigorous structure chosen for the games. The games themselves merely focus on the exercise offered. The students can replay a game, if they need more practice, or can skip a game, if they need less practice.
- **Automation of selected part-tasks**. The games do not focus on automation of selected part-task of the research cycle. At the beginner level aimed for, it was seen as not desirable to focus on any part-task in particular. The games do, however, each follow the four main parts of the research cycle. Games of the same kind repeat the same issues such as identifying the key variables for the study from the scenario, select the design (experimental or correlational), formulate the null hypothesis and interpret the statistical test for the quantitative games or select the appropriate methods and samples and carry some qualitative coding for the qualitative games.

2.2. Game Details

The Quantitative games

In designing the details of the games (see fig 1 and 2) a careful balance had to be maintained between providing the students with enough information and giving them too much information. The games were based on a series of examples where similar questions had to be asked for each example and players had to make decisions for that example. Each example was introduced via a scenario which provided a brief description of a specific research question, participants and variables and the player is guided through the sequence of activities relating to the varied issues which need to be considered in relation to that hypothesis. The decisions were interrelated and had to be considered in parallel. There was not necessarily a correct order in which the decisions were made, although the games led players through the issues in a specific order and players had to make decisions. The intention was that through repetition of the sequence of operations players will pick up the issues which need to be considered. The sequence of activities in the quantitative game was as follows:

- The player reads the scenario
- The player decides whether the design for that study would be experimental or correlational
- The player formulates the null hypothesis for that study. This is implemented via a drag and drop exercise.
- The player identifies the variables (fig 1) and level of measurement of these variables. This is implemented via a hangman game mechanism.
- The player identifies the correct raw data set from a choice of two
- The player identifies correct data summaries (tables and graphical representations)
- The player identifies and interprets the correct statistical test

The two latter stages are implemented by a series of questions structured into a tic-tac-toe (fig 2) game mechanics.

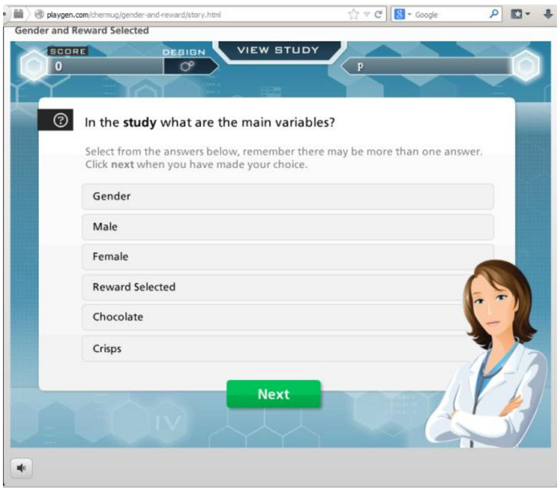

Figure 1. An example of identifying the main variables of a study.

Figure 2. An example of the tic-tac-toe game mechanics.

The Qualitative games

The qualitative games are a set of connected games, the student has the challenge of being accepted in a research group. Level 1 represents how the student prepares for an "aptitude test" for a job in a Research Group for Weight Studies; Level 2 represents the "aptitude test" itself where a qualitative study is simulated assisting the student in that process; in Level 3, the student has to demonstrate the previously acquired skills in working together with the research group team members. The game gives immediate feedback in level 1, general feedback in level 2 and at level 3 only feedback at completion of the game.

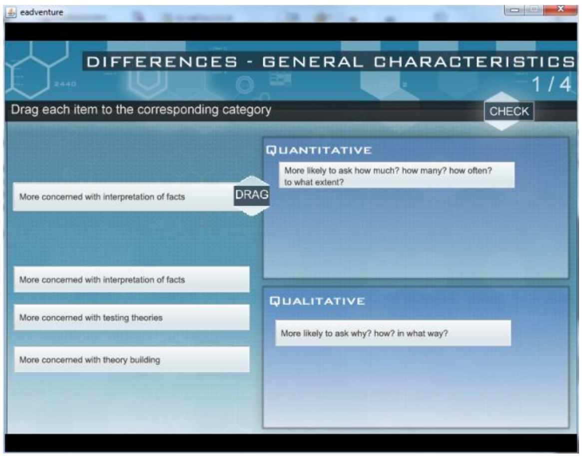

Figure 3. An example of differences between the qualitative and the quantitative approaches.

In level 1 the student has to learn the main differences between quantitative and qualitative studies emphasizing the specific features of qualitative analysis. This level presents exercises where the student learns to distinguish between qualitative and quantitative datasets, to identify differences in terms of theoretical underpinnings, methods, the kind of data collected and the kind of analysis. The feedback provided by the game in level 1 is immediate: the game informs about the correctness of the responses after each attempt and does not allow to continue until the student answer each question correctly.

At level 2 students are presented with a specific research question and background to the study and they are supported as they carry out a sequence of activities which are typical of the qualitative approach. The objective is to show the player the steps and methods to be followed in qualitative research by allowing them to experience these. This level focuses on issues in sampling and data analysis. Sampling is an important issue in qualitative analysis and the sampling activity was designed to reflect the idea that in selecting the sample for a qualitative study careful consideration needs to be given to the size, quality and representativeness of the sample. The sampling part presents different options allowing the students to design their experiment:

- Selecting the methods to gather data
- Deciding how and where the data will be collected
- Deciding the number of people who will participate in the study and their characteristics

The qualitative coding exercise is a key activity with respect to qualitative analysis. In this activity players are provided with participants' verbal statements about a specific topic (food preferences) and a number of pre-defined categories and the goal for the players is to classify the statements according to the appropriate higher level categories. Figure 4 shows an example of the data ("I personally consider weekend(s) more of the party time for lunch and what I mean by party time is like maybe pizzas, hot dogs, and hamburgers and maybe ribs sometimes") and the higher level categories (the child, the food, parent, and context of time) to which players have to assign data. Players are given eight data items to code and are given feedback about the correctness of their responses.

Figure 4. An example of the data and categories for the coding activity.

Level 3 is similar to level 2 in that players are presented with a research question and background to the study and have to make decisions about sampling and carry out data analysis. This level is more exploratory and game-like in that students are not provided with feedback until they have completed the game. The game uses the narrative of sending a paper to a journal: the author prepares the paper and sends it to the journal, then receives feedback and send the paper back to the journal until the reviewer decides the paper is ready to be published.

3. Evaluation Phase 1

The evaluation was carried out in 3 different phases. The first two phases were formative and aimed to assess the usability of the games and collect suggestions for improvements, the third phase was summative and aimed to assess to what extent the games did achieve their goals. Different cohorts of staff and students were involved in each phase (Boyle, McGregor, & Manea, 2014).

- Phase 1 was the preliminary testing of the initial game prototype and involved a small number serious games experts, research methods experts and teacher trainers. Elements of the games were still changeable at this point.
- Phase 2 was the usability phase and involved nursing and social science staff involved in teaching research methods and students taking research methods modules. Surface elements of the games were still changeable at this point.
- Phase 3 of the evaluation was a more rigorous evaluation designed to establish whether the use of the CHERMUG games engages students and helps them to learn about research methods and statistics.

In this paper given the focus on the design and development we will direct the attention on the phase 1 evaluation. For phase 1 volunteers were recruited from serious games experts, research methods experts and teacher trainers at three of the partner sites. In total 13 persons responded to participate in the evaluation of which 5 serious game experts and 8 research methods experts or teacher trainers.

The main objective of the evaluation was to get qualitative feedback on how well the games supported the intended learning outcomes and to get input on how to improve the games. The evaluation was done in a period of 8 weeks. Starting with the first version of the games to the more or less final version, since the evaluation findings were, if possible, incorporated in the games as they emerged.

At each stage, the participants were asked to play one or more of the games. Following the game play they were either interviewed to respond on the games played or they were asked to fill out a questionnaire with some general demographic information and a set of questions related to usability of the games, including questions about the ease of access, game structure and user interface, the usefulness of the games including questions like “this is a nice way to learn about research methods” and “I really learned something” and to what extent the games are motivating including questions such as “the game play is motivating and challenging” and “I would like to be offered more games alike the ones used to support my learning”. Finally, part of the questions related to the strong and the weak points of the game, suggestions and errors and their overall impression of the games.

4. Results

The findings of the phase 1 evaluation centred around the central questions of importance at this stage i.e. usability, usefulness, motivating, strong and weak points of the games and any errors, technical or content wise. During the 8 weeks of the phase 1 evaluation, the feedback, following the improvements implemented, stepwise changed. In the first prototype the participants had sometimes to endure problems due to browser or browser version specific errors or game specific implementation errors. Stepwise the games became more mature and in line with this the focus of the feedback changed from reports on problems identified to opinions on the usefulness of the games and to what extent they were seen as being motivating and could be improved. This aligned well with the scores of the 2 groups that used the questionnaire. Overall, the first group (5 respondents) was slightly positive (average score 2.8 on a five point scale), the second group (2 respondents, 4 weeks later) was much more positive (average score 2.1). Though the overall appreciation was positive many and critical remarks were made. The comments of all 13 respondents are summarised below.

Usability. As mentioned above the first series of comments mainly targeted malfunctioning of the game. At a later stage the comments did largely focus on possible/necessary improvements. A comment returning in the feedback of a substantial part of the evaluators was the need for a more concise introduction to the exercises in the games (“Overall we need more instructions”). The ease of use was too much depending of the skills of the user. One other comment regularly returning was the need for better navigation facilities e.g. being able to go back in the game or at least to get access to the scenario of the game and the choices made and the remark that it would be better to be more consistent between the games in fonts and game graphics.

Usefulness. Many of the comments regarding the usefulness discussed the details of the games at the content level. This ranged from differences in view on the design e.g. some of the questions did not fully fit the research scenario (“At least one or more of the questions were not representative of the research article”) to issues such as whether or not it is essential to have the coding categories to be defined by the students themselves (“you would not have the thematic categories and would get the students to make up the categories") and the comment that if targeting at beginners the knowledge required to play the games should be clear in advance or be part of the games ("Terminology is not always clear for beginners"). Related to this it was commented that on many places more elaborate feedback was desirable. Another comment, in particular important for the possible use of the games in teaching, is the statement that "the game is more about rehearsal, repetition and testing than about instruction". Many of the evaluators expressed the same vision in similar words.

Motivating. In general the games were perceived well ("These games are good ways to teach statistics. They are interactive and the cases are realistic."). However, there were also critical remarks with regard to game experience ("It is more a multimedia exercise as a game"). Additionally, the game experts did classify the games more like instruction with game elements than as games and, related to this, the score mechanism was described as "the rewards are not content related".

Strong and weak points. The strong and weak points mentioned related to the comments given above. Weak points raised were the need to give more instructions or references and more elaborated feedback. Strong points mentioned were the overall obesity case. It seemed relevant, and quite easy for a broad audience. Additionally, the games were perceived as useful to practice and as a formative test.

5. Discussion and Conclusion

Acquiring expertise in research methods and statistics is not easy and poses a significant challenge for many students. The way to teach it and the required subject material, though being a fixed topic in curricula, is in no-way standardised between universities within or between countries. The CTA analysis carried out enabled us to identify the expectations, differences and overlaps in teaching research methods and statistics. Within the differences, it conveniently showed that the experts agreed that all phases of the research cycle are equally important in a first introduction and that both quantitative and qualitative research are seen as important. Together with the results of the literature review and the user requirements it resulted in a set of 11 mini-games which can complement existing methods or be used as refreshment exercises for more advanced students. Notably, the 4C-ID model guided us to set up each game as a complete task covering all aspects of research in one realistic context. The adoption of the 4C-ID model did help us to prepare the global design of the games in an educationally sound way. Since we did not compare the use of 4C-ID with other methods, we can only state that it was a welcome addition to assist in the design of our relatively simple games. The guidelines proposed did easily fit, however, research focussed on this aspect is required to be able to decide to what extent 4C-ID is suitable for the design of various kinds, including more complex, serious games.

Generally, the phase 1 evaluation pointed out that the proposed games were accepted. However, many small and bigger issues were raised. Most of the comments were dealt with either within the games or outside the games e.g. by making supportive materials including a teacher and student guide and clear guidelines and the knowledge required to play the games. The errors revealed were addressed, moreover, a large part of the suggestions was taken into account including e.g. the navigation facilities were improved, the scenario underlying each game was made accessible throughout the games and graphics and fonts were aligned. Some comments remained unaddressed or partly unaddressed, the feedback was improved but the amount remained rather minimal and no references to external learning resources were added. The reason was not to interrupt the flow of the exercises too much. Also the game play remained untouched. The advantage of having a set of small and easy to use simple games was in our perspective more important than having a few much more complex and elaborated games.

Meanwhile, phase 2 and phase 3 of the evaluation have been carried out. The phase 2 evaluation with more than 500 students in three countries confirmed the results of the phase 1 evaluation. The phase 3 evaluation was a more rigorous evaluation designed to establish whether the use of the CHERMUG games engages students and helps them to learn about research methods and statistics. Over 400 students in 4 countries used the games or an alternative. The first results are promising and a detailed analysis of the results is underway. In line with this the partner institutes have made the games available to their students as additional exercise material. Moreover, the games and supportive materials have been publicly released (for more information see www.chermug.eu).

Acknowledgements

The authors would like to thank all CHERMUG staff and the staff who contributed to the design and evaluation. The present work was carried out as part of the CHERMUG project. This project is partially supported by the European Community under the Lifelong Learning Programme project nr. 519023-LLP-1-2011-1-UK-KA3-KA3MP. This document does not represent the opinion of the European Community, and the European Community is not responsible for any use that might be made of its content.

References

[1] Asbell-Clarke, J., Edwards, T., Rowe, E., Larsen, J., Sylvan, E., & Hewitt, J. (2012) Martian Boneyards: Scientific Inquiry in an MMO Game. *International Journal of Game-Based Learning*, **2**(1), 52-76

[2] Boyle, E. A., Connolly, T. M. & Hainey, T. (2011) The role of psychology in understanding the impact of computer games. *Entertainment Computing*, **2**, 69-74.

[3] Boyle, E. A., & MacArthur, E. (2013) User requirements. CHERMUG project deliverable D13. Available online at: http://hdl.handle.net/1820/4847

[4] Boyle. E.A., MacArthur, E., Connolly, T., Manea, M., Karki, A. And Van Rosmalen, P. (submitted) A literature review of games, animations and simulations to teach research methods and statistics.

[5] Boyle, E. A., Manea, M., & Karki, A. (2013) Systematic review of the literature on computer games for education. CHERMUG project deliverable D12. Available online at: http://hdl.handle.net/1820/4846

[6] Boyle, E. A., McGregor, S, & Manea, M., & Van Rosmalen, P. (2014) Evaluation deliverable. CHERMUG project deliverable D21

[7] Boyle, E. A., Van Rosmalen, P., MacArthur, E., Connolly, T., Hainey, T. & et al. (2012) Cognitive Task Analysis (CTA) in the Continuing/ Higher Education Methods Using Games (CHERMUG) Project. In P. Felicia (Ed.), *Proc. of the 6th European Conference on Games Based Learning* (pp. 63-71). October, 4-5, 2012, Cork, Ireland. Academic Publishing International Limited, Reading, UK.

[8] Boyle, E., Van Rosmalen, P., & Manea, M. (2013) Cognitive Task Analysis. CHERMUG Deliverable D14. http://hdl.handle.net/1820/4848

[9] Chipman, S. E. Schraagen, J. M. C. & Shalin, V. L. (2000) Introduction to Cognitive Task Analysis. Chapter 1 in J.M.C. Schraagen, S.E. Chipman, and V.L. Shalin (Eds.) *Cognitive Task Analysis*. Mahwah, NJ: Lawrence Erlbaum Associates.

[10] COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE AND THE COMMITTEE OF THE REGIONS. An Agenda for new skills and jobs: A European contribution towards full employment. (2010) Available online at: http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2010:0682:FIN:EN:PDF

[11] Connolly, T.M., Boyle, E.A., MacArthur, E., Hainey, T., & Boyle, J. (2012) A systematic literature review of empirical evidence on computer games and serious games. *Computers & Education*, **59**(2), 661-686.

[12] De Freitas, S., Rebolledo-Mendez, G., Liarokapis, F. Magoulas, G. & Poulovassilis A. (2010) Learning as immersive experiences: using the four dimensional framework for designing and evaluating immersive learning experiences in a virtual world. *British Journal of Educational Technology*, **41**(1), 69-85

[13] Enfield, J. (2012) Designing an Educational Game with Ten Steps to Complex Learning. (Doctoral Dissertation).

[14] FAS (2006) Harnessing the Power of Video Games for Learning. Summit of educational games. Washington DC, Federation of American Scientists. http://www.fas.org/gamesummit/Resources/Summit%20on%20Educational%20Games.pdf

[15] Halpern, D. F., Millis, K., Graesser, A. C., Butler, H., Forsyth, C., & Cai, Z. (2012). Operation ARA: A computerized learning game that teaches critical thinking and scientific reasoning. *Thinking Skills and Creativity*

[16] Harteveld, C. (2011) *Triadic game design: balancing reality, meaning and play*. London: Springer-Verlag.

[17] Huang, W. D., & Johnson, T. (2009) Instructional Game Design Using Cognitive Load Theory. In R. Ferdig (Ed.), *Handbook of Research on Effective Electronic Gaming in Education* (pp. 1143-1165). Hershey, PA: Information Science Reference. doi:10.4018/978-1-59904-808-6.ch066

[18] Hulshof, C. D., Eysink, T.H.S., & de Jong, T. (2006) The ZAP Project: Designing interactive computer tools for learning psychology. *Innovations in Education & Teaching International*, **43**, 337–351.

[19] Hummel, H.G.K., Van Houcke, J., Nadolski, R.J., Van der Hiele, T., Kurvers, & Löhr, A. (2011) Scripted collaboration in serious gaming for complex learning: Effects of multiple perspectives when acquiring water management skills. *British Journal of Educational Technology*, **42**, No 6, 1029-1041.

[20] Joint Information Systems Committee JISC (2007) Game-based learning: Briefing Paper (online). Available online at: http://www.jisc.ac.uk/media/documents/publications/gamingreportbp.pdf

[21] Keller, J.M. (1983) Motivational design of instruction. In C. M. Reigeluth (Ed.), *Instructional-design theories and models: An overview of their current status*. Hillsdale, NJ: Lawrence Erlbaum Associates

[22] Kirriemuir, J., & McFarlane, A. (2004) Literature Review in Games and Learning, A Graduate School of Education, University of Bristol, published by Futurelab, http://www.futurelab.org.uk.

[23] Klopfler, E., Osterweil, & Salen, K. (2009) *Moving learning games forward; obstacles, opportunities & openness*. Boston MA: MIT - The Education Arcade.Last accessed: 15 September 2011.

[24] Lovett, M. C. (1998) Cognitive Task Analysis in Service of Intelligent Tutoring System Design: A Case Study in Statistics. B.P. Goettl et al. (Eds.): *ITS '98*, LNCS, 1452, pp. 234-243.

[25] Lukosch, Van Bussel & Meijer (2012) A Game Design Framework for Vocational Education in International *Journal of Social and Human Sciences*. Amsterdam **6**(1)

[26] Nadolski, R. J., Hummel H.G., van den Brink, H. J. et al (2008) EMERGO: A methodology and toolkit for developing serious games in higher education. *Simulation & Gaming*, **39** (3), 338-352.

[27] NFER (2009) Teacher Voice Omnibus Survey. Retrieved August 24, 2012, from NFER: Berkshire: http://www.nfer.ac.uk/nfer/what-we-offer/teacher-voice/PDFs/futurelab.pdf

[28] Pernin, J., Michau, F., Mandran, N & Mariais, C. (2012) ScenLRPG, a Board Game for the Collaborative Design of Gbl Scenarios: Qualitative Analysis of an Experiment. *Proc. of the 6th European Conference on Games Based Learning*. Cork, Ireland, 4-5 October 2012, pp. 384-392.

[29] Ramler, I. P. & Chapman, J. L. (2011) Introducing Statistical Research to Undergraduate Mathematical Statistics Students using the Guitar Hero Video Game Series. *Journal of Statistics Education*, **19**(3)

[30] Tishkovskaya, S. and Lancaster, G. A. (2010) Teaching strategies to promote statistical literacy: Review and implementation. *ICOTS8* Contributed Paper.

[31] Van Merriënboer, J. J. G., & Kirschner, P. A. (2012) *Ten steps to complex learning* (2nd Rev. Ed.). New York: Routledge.

[32] Vier in Balans Monitor 2012 (2012) Retrieved December 5, 2012, from: http://www.kennisnet.nl/fileadmin/contentelementen/kennisnet/Over.kennisnet/vier-in-balans-2012.pdf

[33] Williamson, B. (2009) Computer games, schools, and young people. A report for educators on using games for learning. Futurelab: Bristol. Available online at: http://archive.futurelab.org.uk/resources/documents/project_reports/becta/Games_and_Learning_educators_report.pdf