Assessment process for a simulation-based training environment in Global Software Development

Miguel J. Monasor
Lero, University of Limerick
Alarcos Research Group, University of Castilla-La Mancha
MiguelJ.Monasor@gmail.com

Aurora Vizcaíno
Alarcos Research Group, University of Castilla-La Mancha
+34 926 295300 ext. 6478
Aurora.Vizcaino@uclm.es

Mario Piattini
Alarcos Research Group, University of Castilla-La Mancha
+34 926 295300 ext. 3715
Mario.Piattini@uclm.es

John Noll
Lero, The Irish Software Engineering Research Centre, University of Limerick
+353-61-202956
John.Noll@lero.ie

Sarah Beecham
Lero, The Irish Software Engineering Research Centre, University of Limerick
+353-61-233769
Sarah.Beecham@lero.ie

ABSTRACT

Simulation has been applied in several Software Engineering fields, and is shown to be a useful method in industrial training. As part of our research work, we have used simulation to provide training in Global Software Development (GSD). We have developed a platform to strengthen GSD skills by simulating realistic settings in which learners interact with Virtual Agents of differing cultures. Thus, learners will experience multi-cultural problems and will develop specific GSD communication skills.

The development of these skills must, however, be accurately assessed, bearing in mind that the training is aimed at learners with different characteristics and skills. In this paper we present an assessment process based on educational theory adapted to our simulation-based training environment. Methods to minimize the instructors' effort and tailor the assessment to specific training needs are proposed. The assessment process has been evaluated by 34 potential users. Results indicate that the assessment method yields meaningful results and proof of learning. Also that automated assessment can be achieved with minimal intervention from the instructor. Though tailored for GSD, this method could be applied to other domains, in both industry and academia.

Categories and Subject Descriptors

K.3.2 [**Computer and Information Science Education**]: Computer science education;
I.6 [**Simulation and modeling**]: General;

General Terms

Measurement, Design, Experimentation, Human Factors, Languages

Keywords

Assessment, Global Software Development, Training, Education, Simulation

1. INTRODUCTION

In recent years, Global Software Development (GSD) has introduced new challenges in the field of education. In order to develop software in a globally distributed team, practitioners can no longer rely on excellent technical skills. Soft skills such as the ability to collaborate and communicate effectively across a variety of geographic, cultural and linguistic settings has become equally important. Consequently, training in how to cope with cultural and linguistic differences is a recognized necessity [12].

In order to provide training in GSD skills, we have developed VENTURE (Virtual ENvironmenT for commUnication and collaboRativE training) [11], a simulation-based training environment that places learners in a realistic GSD scenario. In this setting, learners can hone required GSD skills in a systematic and guided manner by textually interacting with Virtual Agents (a Virtual Colleague and a Virtual Guide). Virtual Colleagues are characterized by a specific culture and play a role in the context of the problem. A Virtual Guide directs the interactions, corrects the learners and provides learners with feedback in real time. The feasibility of using VENTURE for cultural training purposes in GSD, was evaluated in our previous work [11], with encouraging results.

However, having proven that the concept has potential as a training tool, it is also important to independently assess the learning achieved through using the tool. There are several benefits to combining an assessment with the training. For example, the learners themselves will benefit from the assessment by having a feeling of achievement, and research shows that students are more motivated to learn when there is some 'proof' of learning [14]. Also, in the case of companies, managers will have a better view of the workforce as a whole, and will be able identify where individual strengths and weaknesses lie in communicating with different cultures and across different languages. This knowledge will help identify where further training is needed, and how best to use the available staff resources.

Fulfillment of the requirements to equip students with professional competences, can be hard to demonstrate in practice [7]. A good assessment must be valid, reliable and transparent, and must clearly outline the goalposts [13]. Assessment procedures can vary depending on the type of training, tools employed, training methodology, individuals involved or field of application [13]. Although lessons learned in creating GSD learning environments are reported [6], methods for GSD learning assessment is still an area in need of further research. Daniels [7] discovered that the assessment aspect in an international open-ended group project is strongly influenced by how competencies are supposed to be improved in a learning environment. The research question that guides this study is: "How can we assess GSD competencies in a simulation-based training environment?"

ITICSE '14, June 21 - 25 2014, Uppsala, Sweden
Copyright 2014 ACM 978-1-4503-2833-3/14/06$15.00.
http://dx.doi.org/10.1145/2591708.2591747

The objective of this paper is to present and evaluate the assessment process implemented in VENTURE in the field of GSD. Minimizing the instructors' effort and tailoring the assessment to specific training needs are also considered.

2. RELATED RESEARCH

The effectiveness of educational training environments has been assessed in different ways [9], [10]. Minimizing the instructors' workload is a common requirement in these systems, although not all aspects of assessment can be automated. As an example, a semi-automatic web-based assessment tool has been presented by Heo [9] with the aim to help instructors comment directly on students' submissions, provide students with access to this information, and measure their learning through quizzes.

Peer- and self-assessment is often used to track the learning process where instructors delegate the assessment responsibility to their students [4]. This type of assessment also plays a role in the motivation of the students by giving them goals to achieve and a means to follow their progress.

Assessment in Software Engineering education

Most of the current research in Software Engineering education is focused on the original Bloom's Taxonomy [5]. However Whalley et al. [16] suggests some deficiencies in the Bloom taxonomy when applying it to programming problems. The main difficulty with these approaches is assessing the students' improvement, as the scores are not linked to previous knowledge and skills. Some methods deal with this problem by assessing the students several times during the course [2]. The research in this field, therefore, shows the need to assess students before and after training in order to measure their progress.

Assessment in GSD training

Despite the lack of studies on assessment on the GSD field, Damian et al. [6] report on an experience of a GSD course conducted in collaboration with three universities. The course emphasized requirements management activities through synchronous computer-mediated relationships, incorporating metrics for assessing GSD skills in the fields of [6]: international teamwork, distributed project management, computer mediated communications and ambiguity/uncertainty. Authors applied these metrics (where different categories were weighted according to its importance), to produce a final score.

Assessment in multi-cultural training

In the related field of multicultural training, there are tools oriented towards providing training to multinational corporations and expatriates. Cross Cultural Adaptability Inventory (CCAI) [8] is an online tool that addresses learner's ability to adapt to any culture. By means of a 50 question test, it provides self-assessment covering four areas: emotional resilience, personal autonomy, flexibility/openness and perceptual acuity. A guide helps participants understand the implications of their scores and suggests actions that can help them strengthen their skills.

The Cultural Orientations Framework (COF) [15] is a framework to assess and compare cultures based on the following cultural dimensions: sense of power and responsibility, time management approaches, definitions of identity and purpose, organizational arrangements, notions of territory and boundaries, communication patterns, and modes of thinking.

In summary, the assessment methods reported in the literature are varied, and create a useful set of measures we can apply to our VENTURE training platform. For example, our assessment process should include self-assessment methods to identify learners' improvement, and incorporate automatic assessment that take advantage of VENTURE's simulation capabilities. The next section explains how these assessments are applied in VENTURE.

3. LEARNING ENVIRONMENT

VENTURE [11] is based on an e-learning platform in which instructors can set up scenarios, upload teaching material, communicate with learners and monitor their progress. Learners can study theory online or/and attend classroom based classes, which are not always an option in industrial based teaching. A Chat Simulator (see Figure 1) allows learners to train on synchronous interactions. The learner's objective is to obtain as much information as possible from the Virtual Colleague within the time limit of the simulation. The Virtual Guide will provide learners with instant feedback during the interaction.

Figure 1. Chat simulator.

VENTURE's training objectives are to:

- provide knowledge of theory and best practices of GSD and teamwork in multicultural environments;
- expose learners to the multi-dimensional and multi-cultural nature of GSD through simulation;
- help students to acquire GSD skills and more specifically, cultural and linguistic skills;
- minimize the learners' and instructors' effort and coordination problems; and
- encourage students to engage through self- assessment.

4. ASSESSMENT OBJECTIVE

A key challenge for the VENTURE training platform is to objectively assess the level of learning achieved through this independent style of training. The aim of the assessment is twofold: to provide learners with advice on the skills they have and those they must improve, and to help the instructors to create training scenarios based on each learner's experience. More specific objectives are to:

O1. Identify skills in which the learner needs further training.
O2. Assess the learner's ability to put into practice or retrieve knowledge learned in training sessions.
O3. Monitor the learners' progress (knowledge and skills).
O4. Use unambiguous, systematic, reliable and objective assessment criteria.
O5. Promote knowledge construction.
O6. Minimize time and resources needed.

5. ASSESSMENT PROCESS

In order to achieve the objectives listed in Section 3, we have designed an assessment process for the simulated learning environment. This assessment incorporates concepts taken from assessment strategies found in the related literature, as discussed in Section 3. These fall into four kinds of assessment, presented here.

1. **Diagnostic assessment** determines a learner's starting level of knowledge and skills. The educator will determine the assignments for the learners based on this initial level. The scores earned in this assessment do not directly contribute towards the final grade of the learners, but provide a base-line measurement to act as proof of their advance at the end of the course or unit.

2. **Formative assessment** by means of simulation, learners can experiment with different ways to communicate, through feedback, asking questions and taking risks that they would not necessarily take in real settings. VENTURE provides feedback and direction to the learners so that they can improve their communication skills, which are assessed in real time.

3. **Summative assessment** measures the achievement relative to the course objectives; in our case the evaluation of artefacts and exams.

4. **Self assessment** by means of surveys and questionnaires is also used to allow learners to judge their own skills.

In order to achieve the Course Learning Objectives (CLOs), learners are assigned reflective tasks, learning from their successes and failures through provision of immediate feedback. The system gathers these reflections by automatically monitoring their advances and measuring their skills. Moreover, instructors can receive feedback from students and examine their self-assessments. The analysis of these reflections is used to examine whether learners achieved their goals. During the simulations, reflection is promoted by: explaining the consequences of certain actions, proposing alternatives or giving the learner the opportunity for finding alternatives, allowing learners to see the problem from various perspectives; and placing learners into different and varied contexts.

Activities	Objective	Assessment	Detail
A) Initial learners´ categorization	Determine the training that best fits with each learner	Identify learners´ personal details and current skills by means of self-assessment	Online questionnaire in Appendix A. Interviews
B) Course preparation and customization to learners´ needs	Adapt materials to learners´ needs. Assignment of training modules and schedule	Preparation of adapted GSD rules. Inform learners about assessment objectives	Set the severity of GSD rules. Establish weights for the training methods
C) Study the theoretical material	Provide theoretical knowledge on GSD	Assessment of theoretical knowledge	Questionnaires, quizzes or exams
D) Execute training simulations	Provide practical experience in GSD	Assessment of interaction skills. Instructor provides feedback and adapts the learning	Reports after simulations. Logs of simulations. Evaluation of deliverables
E) Summative assessment	Gather assessment information	Learners´ self-assessment , final evaluation and assessment of the learners' progress	Final evaluation considering final reports. Questionnaires in Appendix A and B.

Figure 2. Assessment process.

The assessment process developed (shown in Figure 2), is described according to the following activities of the training process:

A) Initial learners' categorization (*diagnostic assessment*): Learners are assigned an initial predefined baseline questionnaire for identifying their current skills. Based on an online questionnaire, learners' culture and skills are determined. With these results, instructors can decide the training that best fits each learner's needs based on their skills and characteristics.

B) Course preparation and customization to learners' needs: In a further step, the kind of training that best fits each learner is determined and defined by the instructor. Instructors prepare the theoretical and training materials based on this categorization. Training modules are designed focusing on cultural dimensions and linguistic problems suited to each learner.

In case of specific needs, instructors can also adapt existing scenarios by customizing cultural, linguistic and GSD rules according to their individual training objectives.

C) Study the theoretical material: In VENTURE, the Teaching and Learning Activities (TLAs) consist of theoretical lessons and practical simulations. The cycle for learning theory and executing practical simulations is repeated with each lesson until completing the course scope. For each assignment learners must execute the associated training simulations (by chat or email) and submit their associated Assessment Tasks (ATs), basically answers to questionnaires. The information of their evaluation along with the information provided by the automatic reports generated by the simulators, will also serve to assess the learner (*formative assessment*).

D) Execute training simulations: After completing the different lessons, learners complete a questionnaire regarding concepts learned and they execute the training simulations. The instructor will monitor the accomplishment of each learning objective. This is achieved by reviewing the logs of the learner interactions with the chat and email simulators, which are stored in the system. These logs provide formatted and organized information about the rules that were fired and the context in which they were fired including the students' interaction. At the end of the simulation, prior to producing the report, the students can review any mistakes committed, one by one. After each lesson, learners receive the instructor's comments and future lessons can be adapted accordingly.

E) Proof of learning: The Course Intended Learning Outcomes (CILOs) are measured by means of a *summative assessment* which encompasses the different marks of the final reports of the simulations, questionnaires and exams. Each one of these ATs can be assessed by using different weights in the final score, according to the instructor's criteria.

Moreover, after completing each simulation, learners revisit and answer the same questionnaire that they answered at the start of the simulation. In this way GSD skills are quantified and compared with the results of the benchmark questionnaire. Learners are given an additional proof of learning questionnaire to quantify their improvement through measuring their perception on a 1-5 scale (*self-assessment*).

5.1 Self-assessment

Self-evaluation through reflection is promoted by means of questionnaires. Automating the self evaluation process provides learners and educators with a valuable mechanism for feedback, and a means to measure course progress online [3].

As VENTURE's design is aimed to minimize the training and assessment effort, the automation and ease of use of the assessment procedure is essential for its application in both universities and companies.

The first activity of the assessment process includes a self-assessment questionnaire where knowledge is checked relating to key aspects of effective work in GSD settings. In a similar vein, our initial baseline questionnaire will serve to categorize the learners at the first stage of the assessment process by considering not only multiculturalism, but also linguistic and GSD skills. This questionnaire gathers information about how well learners can perform in GSD skills. Learners will answer the questions in a 1-5 scale.

At a final stage of the process, after course completion, the questions answered at the start of the assessment are revisited. The objective is to evaluate the learners improvement at the end of the training process by comparing their knowledge at the beginning and at the end of the course.

Finally, at the end of the course, learners also fill in a proof of learning questionnaire, which serves to quantify how much they improved their GSD skills. The difference with the previous questionnaire is that in this case, learners are directly asked about their perceived improvement.

5.2 Automated assessment

The training scenarios comprise several cultural, linguistic and GSD rules. As an example, we can have cultural rules that fall into several categories such as use of direct/indirect speech, use of formal/informal language, uncertainty avoidance, etc. Each learner starts with a perfect score of 100, where, if no errors are made, throughout the interactions, their final score will be 100. However, each rule contained in the scenario is also defined by a severity mark, indicating the score that will be subtracted if the user fires such a rule. Therefore, the simulator can automatically assess the learners' performance when interacting with Virtual Agents. After the execution of a training scenario, an automatic assessment is presented in a report detailing learners' performance in the different areas trained (see Figure 3).

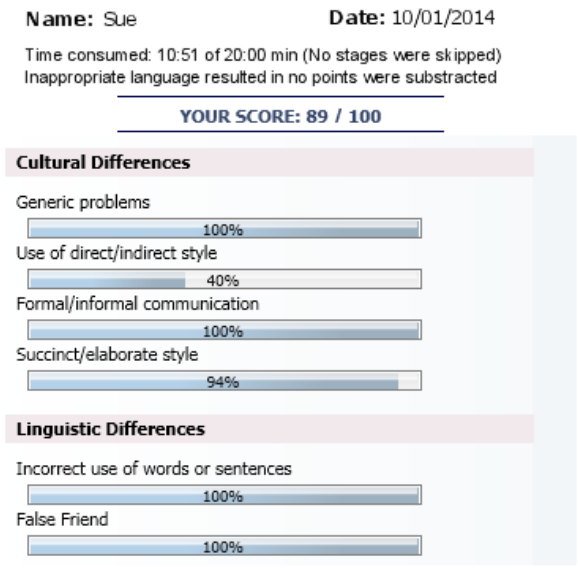

Figure 3. Automatically generated final report.

The final report includes information about the date and duration of the simulation, as well as the number of times that the user gets stuck during the simulation and has to skip a stage. The report also includes a final score and sectional results, detailing one by one the score obtained in each GSD skill that has been trained in the scenario. Use of inappropriate or offensive language during the simulation can also take points away of the final score.

The scores for each skill are calculated by taking into account the total number of learner's mistakes of a specific skill, considering the severity of the rules fired during the simulation.

By analyzing this report, learners will gain an insight of their progress. They can also examine the log of the conversation, in which the Virtual Guide provides them with feedback. Both, the information contained in the report and log of the conversation are stored in the database, so can also be inspected by the instructors.

6. EVALUATION

We conducted a field study to evaluate the assessment component of VENTURE. The objective was to determine whether the assessment effectively monitored the students' progress relating to the cultural, linguistic and GSD-related problems simulated by VENTURE.

A within-subjects design [1] was applied, in which comparisons are made between two or more results obtained from different circumstances involving the same participant. The experiment was conducted by individual participants in two sessions of up to an hour each. These sessions tested the same configuration of the tool with different training scenarios. The CLOs focused on improving participants' performance when interacting with people from Mexico (Scenario A) and from India (Scenario B). Participants played the role of onsite coordinators interacting with an onsite coordinator from Mexico and India respectively. Both training scenarios (Scenario A and Scenario B) provided training on similar skills, so that it was possible to compare the learning effects. The CILOs are related to raise cultural and linguistic awareness, use of appropriate communication styles, ways to avoid misunderstandings, conflicts resolution and negotiation.

Thirty-four participants representing ten different nationalities completed two scenarios and the associated surveys. The participants were potential users of VENTURE comprising 16 students, 12 researchers and six practitioners/researchers. Students represent the largest group and reflect the need to test VENTURE on the main target group. The following questionnaires were applied:

- *Baseline Questionnaire:* designed to collect the participant's demographic information and cultural background.
- *Pre Training Questionnaire*: including questions on the specific knowledge of what is taught on the course (i.e., their current knowledge of GSD issues; confidence levels in communicating with people from other cultures in the job).
- *Post Training Questionnaire*: whose objective was to check whether there was an improvement in the participant's knowledge. This included the same questions as the Pre Training questionnaire, but also gathered the participants' perception of their improvement.
- *Opinion Questionnaire*: this was conducted at the end of the experimental sessions and obtained information on how the students felt about their learning and assessment.

The following data was collected through the evaluation process:

- *Participant´s opinion* of their learning and general experience with VENTURE (*Opinion Questionnaire*).
- *Participant's performance* via VENTURE's automatically generated logs and reports – where all mistakes logged.

- *Participant´s learning* - a comparison of student's knowledge before and after training through pre and post questionnaires and comparative analysis of the automatic logs and reports generated in the different training sessions.

6.1 Procedure

The participants were split into two groups: G1 and G2. Both groups conducted the same tasks (as embedded in Scenario A and B) but in a different order (Evaluation G1, and Evaluation G2).

Both groups filled in the *Baseline questionnaire* as the first task. G1 then participated in Evaluation G1. Evaluation G1 executed ***Scenario A*** first followed by ***Scenario B***. G1 filled in a *Pre Training Questionnaire* for *Scenario A*, after which *Scenario A* was executed, and finally a *Post Training Questionnaire* for *Scenario A* was filled in. After this, G1 participated in *Session B*, which consisted of filling in a *Pre Training Questionnaire* related to *Scenario B*, after which *Scenario B* was executed, and finally a *Post Training Questionnaire* on *Scenario B* was filled in.

The order of the sessions was different for G2, who participated in Evaluation G2, signifying that they executed ***Scenario B*** first, followed by ***Scenario A***. After completing both sessions, both groups of participants (G1 and G2) completed the *Opinion Questionnaire*.

The reason for changing the order in which the two groups (G1 and G2) participated in the scenarios was to test whether there was any bias in the difficulty ratings of the two scenarios. For example, if all the participants who participated in Scenario A followed by Scenario B showed improvements, this might have been because Scenario A was more difficult than Scenario B; however, if Scenario A was always more difficult for the participant (independent of the order), it would not be possible to *compare* the two scenarios for any proof of learning.

6.2 Results

At an initial stage of the analysis it was observed that students received better automatic assessments scores in one of the scenarios than in the other regardless the execution order. After applying the Mann-Whitney U Test the values were obtained ($U = 265.5$, $p=<0.001$) showed that it is highly likely that this difference is not a random occurrence.

These results indicate that one scenario is more difficult than the other (as shown by consistently lower scores). This bias must be taken into account for the subsequent statistical analysis, as it will not be meaningful, from a statistical point of view, to compare all the automatic assessments across the two scenarios. However, each scenario includes the same number of rules of each type and on close inspection we observe that one of these rules was triggered in a similar way in each scenario. As there is no significant difference between scenarios (for this rule) it is possible to test whether any learning has occurred from one scenario to another using this dimension.

Based on this rule, the scores from the automatic assessment method served to detect an improvement in the participants' performance when they execute the second scenario (which can be either scenario A or B, depending on which order they executed the scenarios), after the experience of having executed the first one.

These results signify that the automated assessment is able to provide accurate information for drawing conclusions that are helpful for the instructors. These results also signify that the assessment can be used to compare the performance among the different students using the *same scenarios* instead of comparing the performance of a student in different scenarios even if they are similarly designed.

By analysing the results from the pre and post questionnaires for each scenario, we observed that in both cases there is an improvement in the mean and the median of the answers provided. These differences were shown to be statistically significant by means of the Wilcoxon test applied to the pre and post questionnaires. Learning was observed in the three areas of knowledge tested: cultural differences, linguistic differences and understanding of GSD problems and concepts.

In order to check the accuracy of the assessment method, the participants' perceptions of their learning were requested at the end of the course by means of a new questionnaire. The answers served to validate that participant's perceptions were aligned with the results obtained with the automated assessment method.

6.2.1 Participant's perception

Analysis of responses given in the *Opinion Questionnaire* revealed that participants' perception about the use of VENTURE was positive, who found the training helpful, engaging and usable. The following questions were answered by the participants through a 1 to 5 likert scale: 1) The feedback provided facilitated learning, 2) I was aware of the learning outcomes at the beginning of the course, 3) The assessment information is easy to understand and interpret, 4) The reports helped me to understand the topics trained, 5) The scores I got seem fair, and 6) The assessment motivated me to improve my performance.

The analysis of the results shows that the median of the answers for each one of those questions was 4 out of 5, obtaining also an equal value for the mode, which is quite encouraging. Figure 4 shows the distribution of the answers for the first question, showing that students and researchers (main target users) are the most positive groups.

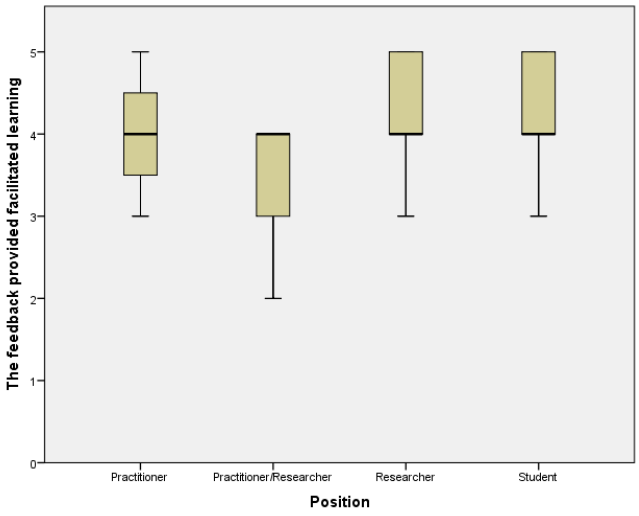

Figure 4. Perceptions of assessment method.

In addition to the quantitative analysis previously analyzed, the surveys also gave the participants the option of providing their opinion. The concept of a simulation-based platform with automated assessment was well received. Finally, after the analysis of the open ended responses of the *Opinion Questionnaire*, the inclusion of additional tutorials at the end of the simulation was suggested as a means to reinforce those points with which the student had more problems. In this respect, the inclusion of additional introductory material on the different

cultures before initiating the simulation was also suggested as a means of reinforcing the process.

7. SUMMARY AND CONCLUSIONS

This paper presents a proposal of a systematic assessment process for a simulation-based training environment. The assessment process is shown to provide both the student and instructor with realistic measures of learning. Advantages of the assessment process are aligned with the objectives mentioned in Section 4:

- Guided adaptation to specific learning needs by automatically assessing the different GSD competences (O1).
- Provision of reliable and systematic assessment based on objective *rules* (O2, O4).
- Monitor learners' progress through logs and reports, minimizing time and resources (O3, O6).
- Promote knowledge construction by providing accurate feedback and assessment to the students (through Virtual Guide and reports) (O5).
- Independent training and minimum intervention from the instructors through the integration with VENTURE (O6).

The evaluation of the assessment process generated the following outcomes:

- The outcomes of the assessment were well valued by the participants.
- The interactions with the Virtual Agents were considered acceptable by the participants.
- Participants improved their knowledge and recognized having learnt from the experience.

The lessons learned will be used to increase the automation of the assessment without compromising the validity of the process as a whole. The information required for the analysis of the assessment was easy to access through the reports automatically generated by VENTURE. This indicates that the time and resources needed by instructors to validate VENTURE related assessments will be similarly reduced. Also the assessment and associated training activities can be adapted to the specific needs of learners, and to instructors' objectives. Moreover, the information gathered during the assessment process can be used to analyse where improvements can be made in the training scenarios.

VENTURE will be shortly applied in a Spanish university serving as support for a subject on Virtual Teams, and companies are also being approached with the objective of refining the assessment process. Future plans are to enhance the assessment reports to include tutorials that reinforce those points shown to cause problems to the learner. Ensuring students' engagement and improving learning outcomes are likely to be key success factors in the take up and effectiveness of the training platform.

8. ACKNOWLEDGMENTS

This work was supported, in part, by Science Foundation Ireland grant 10/CE/I1855 to Lero - the Irish Software Engineering Research Centre (www.lero.ie). It has also been funded by the GEODAS-BC project (Ministerio de Economía y Competitividad and Fondo Europeo de Desarrollo Regional FEDER, TIN2012-37493-C03-01). It is also supported by GLOBALIA (PEII11-0291-5274), Consejería de Educación y Ciencia, Junta de Comunidades de Castilla-La Mancha.

9. REFERENCES

1. Akar, E., Öztürk, E., Tunçer, B. and Wiethoff, M. Evaluation of a collaborative virtual learning environment. *Education + Training*, *46* (6/7). 2004, 343 - 352.
2. Alaoutinen, S. and Smolander, K. Student self-assessment in a programming course using bloom's revised taxonomy *Proceedings of the fifteenth annual conference on Innovation and technology in computer science education*, ACM, Bilkent, Ankara, Turkey, 2010, 155-159.
3. Barzilay, O., Hazzan, O. and Yehudai, A. Evaluation of a software engineering course by reflection. *SIGCSE Bull.*, *41* (3). 2009, 273-277.
4. Black, P., Harrison, C., Lee, C., Marshall, B. and Wiliam, D. *Assessment for Learning: Putting it into Practice*. Open University Press, 2003.
5. Bloom, B.S. *Taxonomy of Educational Objectives Book 1: Cognitive Domain*. Addison-Wesley, 1984.
6. Damian, D., Hadwin, A. and Al-Ani, B., Instructional design and assessment strategies for teaching global software development: a framework. in *International Conference on Software Engineering (ICSE)*, (2006), ACM Press New York, NY, USA, 685-690.
7. Daniels, M. Developing and Assessing Professional Competencies: A Pipe Dream? ; Experiences from an Open-ended Group Project Learning Environment, Uppsala University, 2011, 107.
8. Davis, S.L. and Finney, S.J. Examining the psychometric properties of the Cross-Cultural Adaptability Inventory. *Educational and Psychological Measurement*, *66*. 2006, 318-330.
9. Heo, M. A learning and assessment tool for web-based distributed education *Proceedings of the 4th conference on Information technology curriculum*, ACM, Lafayette, Indiana, USA, 2003, 151 - 154.
10. Huizinga, D.M. Identifying topics for instructional improvement through on-line tracking of programming assignments *Proceedings of the 6th annual conference on Innovation and technology in computer science education*, ACM, Canterbury, United Kingdom, 2001, 129-132.
11. Monasor, M.J., Vizcaíno, A. and Piattini, M. Cultural and linguistic problems in GSD: a simulator to train engineers in these issues. *Journal of Software Maintenance and Evolution: Research and Practice (Special Issue on Global Software Engineering)*, *24* (6). 2011, 707-717.
12. Monasor, M.J., Vizcaíno, A., Piattini, M. and Caballero, I. Preparing students and engineers for Global Software Development: A Systematic Review *International Conference on Global Software Development (ICGSE)*, IEEE Computer Society, Princeton, NJ, USA, 2010, 177-186.
13. O'Neill, G., Huntley-Moore, S. and Race, P. *Case Studies of Good Practices in Assessment of Student Learning in Higher Education*. AISHE, 2007.
14. Olina, Z. and Sullivan, H. Student self-evaluation, teacher evaluation, and learner performance. *Educational Technology Research and Development*, *52* (3). 2004, 5-22.
15. Rosinski, P. *Coaching Across Cultures: New Tools for Leveraging National, Corporate, and Professional Differences*. Nicholas Brealey Publishing, 2003.
16. Whalley, J.L., Lister, R., Thompson, E., Clear, T., Robbins, P., Kumar, P.K.A. and Prasad, C. An Australasian study of reading and comprehension skills in novice programmers, using the bloom and SOLO taxonomies *Proceedings of the 8th Australasian Conference on Computing Education - Volume 52*, Australian Computer Society, Inc., Hobart, Australia, 2006.