The Implementation of Team Based Assessment In Serious Games

Laz Allen and Matt Seeney
TPLD
Dundee, Scotland
laz@tpld.net | matt@tpld.net

Liz Boyle and Fiona Hancock
School of Social Sciences
University of the West of Scotland
Paisley, Scotland
liz.boyle@uws.ac.uk | fiona.hancock@uws.ac.uk

Abstract — **Serious games differ from commercial games with their focus on learning and training and their use of pedagogies. Serious games have several components that can help to make them successful – learning theories, high-quality gaming experiences, offering collaboration and competition as well as the opportunity for assessment. Team assessment has a potentially major role to play in serious games, allowing individuals to evaluate the role they play in a team, what team they fit into best and how their teams operate as a whole.**

The current study set out to examine the potential of the Infiniteams Island game (TPLD) as a learning tool, using a pre-post test method. As Reference [1] proposed, the user would potentially "over estimate" their abilities pre game, with a more accurate response to their skills post game, having a greater understanding of what was being examined and being able to see the skills in action. Using two hundred and four psychology students from the University of the West of Scotland (UWS), each gave answers to a pre-questionnaire, played the Infiniteams Island game and completed a post-game questionnaire.

This study has led the way for future research in the area of team assessment in serious games, such as this one. Having found that team role assessments such as the Belbin Team-Role Inventory cannot be automatic assessments implemented into open-ended free choice games, this study established a basis not only for future use of team assessment, but also demonstrated the capabilities for such serious games to cause players to evaluate their team skills more accurately, but indirectly. Future implementation of team assessment, whether it is team roles or team strengths, should be applied externally, in parallel to any game.

Keywords-team-based, assessment, methodology, technology, cross-disciplinary.

I. INTRODUCTION

Serious games differ from commercial games primarily based on their function. Serious games are entertaining, but are primarily designed with education and training in mind, whereas commercial games focus solely on entertainment. Additionally, the serious game market utilizes not only experienced gamers, but also new players, and incorporates a wide market from the military to health professionals and the government to educational establishments [2]. Consequently these games have to be good quality games and be more readily available than commercial games. Another variation that separates serious and commercial games is the implementation of pedagogy, and that serious games have "more than just story, art, and software", where [3] defines a pedagogy as "activities that educate or instruct, thereby imparting knowledge or skill" and it this that makes games serious. However the pedagogy must follow on from the entertainment aspect (story and art), and not be a separate entity.

Reference [4] refers to serious games focusing on "leveraging the power of computer games to captivate and engage end-users for a specific purpose". Such a usage allows for the user to experience conditions that would otherwise be costly, unsafe and/or time consuming and can have positive impact effects of their behaviour, attitudes and development of skills [2].

Reference [5] stated that "serious games that act more like utilities and exist beyond education offer a lot of promise for the field's future", and he discussed in Gamasutra [6] about how people adopt different definitions and understandings about what serious games actually are, and what they are about, which, in the long run can lead to an ill-defined bigger picture. Reference [6], however, did go on to discuss how all games, not just serious games (although they do play a key role), are changing as a medium. This change is good for the industry, although to keep up with this change there is a need to "broaden" the gaming world, as well as keep producing fresh ideas to improve game play, and to provide "new reasons to game".

According to [7], using components such as, but not limited to game universes over training (an immersive environment to be experienced and explored by the player, for them to experiment and ruminate), seamless play and learning (learning outcomes and problems faced through game play needs to overlap with the users learning and knowledge acquisition experiences), as well as subject and content (this refers to content credibility where the game is created in collaboration with subject experts, the end-user and researchers), the user will remember the content from the game(s) because the game will captivate and motivate the learner to the extent that the user will want to remember (and possibly re-experience) the content they have learned, through this method.

978-0-7695-3588-3/09 $25.00 © 2009 IEEE
DOI 10.1109/VS-GAMES.2009.12

The casual games market has been growing at a rate of twenty percent annually [8], and the serious games market is "turning into a fast-growth industry" [9], so "as games continue to displace television as a mainstream leisure activity, there has never been a better time to study games and to create solid connections between game developers and academic researchers" however, the only problem appears to be the lack of common ground [10]. Reference [10] justifies combining the forces of game developers and academia by discussing the easy nature of recording a player, and the ease at which the games can be replayed and modded, as well as the possibilities for multiplayer cooperation, collaboration and competition. The collaboration of game developers and academics is a step forward for the gaming world, with collaboration on issues such as technical difficulties, and psychometric assessment as well as opinions of domain experts.

It is also possible that educationalists and psychologists could work together to improve the vast domain of computer games, namely due to the educational opportunities that games can provide. Educationalists and psychologists can have varying views on an individual's learning experience, but by working together they can combine their views on computer games and enhance the learning experience. The differences between the educationalist's and psychologist's views on computer games lies primarily on the focus of their research: the psychologist has a tendency to focus on the anti-social and violent aspects of games whereas educationalists like to examine the potential, positive learning effects a game can have [11]. In this respect the "educational perspective and psychological perspective are worlds apart" [11], even though they both look at what can be learnt from games. The next step would be for educationalists and psychologists to work together to focus their research to begin a process that would improve learning. That said, educators and psychologists both believe that students will learn most effectively when they have to use their own abilities and use the processes of inquiry for learning content, over the monotonous task of memorization [12, 13].

In addition, there is the viewpoint that there is a need for co-operation between game developers and psychologists, and [14] is not alone in thinking this, as [15] supports this view. There has been a lot of development in the past decade with this co-operation. In 1999, it was the worry of [16] as to why "game theorists have not turned more frequently to psychologists for information about the learning and information processing processes used by humans", but according to [15] psychologists play a more active role in game research. There are, for [15], three main perspectives for game research: game-rules, game-world and game-play – and it is the latter area which psychologists are involved in researching. However, [15] does state that game developers and psychologists "have no overlap in terms of methodologies."

The methodology used to create a serious game is important– without the right methodology the game will not stand up against its competitors and will probably flop in the market. There are several approaches to the methodology of serious game creation. Learning theories, drawing on the field of psychology, have a key place within serious games, and a "failure to base serious game design on well-established learning theories… increases the risk of the game failing to meet its intended educational goals, yielding a player base who is entertained but who have not acquired new skills or knowledge" [17]. Possible learning theories that can be incorporated into games include Keller's ARCS model (game design), Gagne's Nine Instructional Events (natural application in the structure of the game), Bloom's taxonomy (skill transfer and retention) [17] as well as the ADDIE model [18].

Another methodology, suggested by [19], is to reduce complexity. According to [19] there are two types of game complexity: designed complexity and emergent complexity, and although complexity is a "keyword for serious games" it should be controlled. What [19] wanted to do was to reduce designed complexity: controlled by the developer, and due to the nature of learning through education, this field is "prompted to create complexity by design" [19]. Reference [19] proposed three levels at which design complexity could be reduced: conceptual, practical and technical. Conceptual complexity refers to game dynamics and game locations – including objects and subjects as they allow for communication, collaboration and competition. Practical complexity refers to feedback, where complexity would be reduced using an overall "strategic performance feedback" and by utilizing peer feedback; game structure, where the game developers would focus on "option width rather than option depth" and game representation – where content credibility would reduce complexities. Technical complexity refers to the implementation of a game authoring tool system to simplify the game creation process, which would "follow directly from the conceptual framework".). To demonstrate the effectiveness of a tooling system (on a practical level), the researchers evaluated the use of an authoring tool by eight teachers. All teachers could use the tool without any assistance, and the teachers commented on the tools "flexibility and the large range of learning scenarios it caters for".

Reference [20] proposed that Agile programming is another productive approach to game creation, that would be "effective and suitable for many situations and environments" where it places emphasis on interactions, collaboration with the customer and promotes change. For [20] the Agile approach means that software development is an incremental, straightforward, adaptive and cooperative process, and "Agile thinking is a people-centric view to software development…provide a novel way of approaching software engineering problems" however, scalability can be an issue. Reference [21] also suggested the Agile approach has the capabilities to: balance flexibility and structure; create and respond to change; draw creativity and innovation out of a development team; and lead organisations through turbulence and uncertainty. Reference [22] puts forward a few important points about Agile development. First is the use of product backlogs – these are lists of what is needed, with details of client requirements and any notes needed. The next stage of Agile development is sprint meetings: this is a critical meeting that enables the team of developers to work unaided for up to a few weeks, and goals are defined at this stage. At each sprint the most important product backlogs are taken and added to the "to do" goals list. For Agile software development there has to be a focus on collaboration and interaction, and adapting and

responding to changes. For Agile development to work, there has to be daily communications about the goals achieved and what still needs to be achieved, with a frequent delivery of the working software to the client, because even small successes that lead to functional software is a measure of progress on a project. Agile software development is often unplanned and flexible with outcomes frequently measured.

Another "key challenge for serious games designers" is to "find an optimal combination of delightful play and achieving specified learning outcomes" [18], which relates to [19]'s need for a tooling system to reduce complexity. To attempt to combat this issue [18] developed EMERGO. EMERGO's main goal is to deliver a methodology and generic toolkit for developers of serious games (where these games' key goal is focused on the acquirement of "complex cognitive skills in higher education". The EMERGO methodology employs the ADDIE approach – again another learning theory – as its base. The ADDIE model is a five-phase model and deals with Analysis, Design, Development, Implementation and Evaluation. In addition to this, EMERGO also makes use of Unified Processing, which cycles the ADDIE process, instead of using a waterfall process. From initial studies, [18] found that the EMERGO toolkit was successful in supporting the developers of serious games when it comes to functional and resourceful gaming experiences. The researchers also found, however, that although the intended aim of EMERGO was to focus on complex cognitive skills, it was also efficient when targeting more simple skills, as well as being a useful tool out with the higher educational field. EMERGO provides a platform for semi-expert teams to come together and create efficient games.

Assessment is a wide field, even within the serious games industry with different companies developing their own techniques and methods of assessment. "Assessment of user-player experience in serious games provides a means to validate learning and design" [23], and it has been shown that assessment is crucial [24] and assessment of "content is *the* key component" in learning, especially online. However there are some difficulties in creating excellent assessment engines, such as the process being lengthy and costly [25]. Assessment can also be made difficult because there is not the same need for rote memorization. There are abstract skills, which are difficult to measure, and there is not the same definition of cheating within games as there is in learning [24].

Assessing team skills is no different. There are numerous ways of assessing, or researching, self-perceived evaluations of team skills, including but not limited to the Acumen Team Skills assessment (Acumen), Profiles team analysis (Galliard), the Acumen Team Workstyles (Acumen), the Belbin Team Role Profiling assessment (Belbin) and the Navigator Team Effectiveness tool (Insights). These can be administered in a way to suit the researcher – electronically or paper-based.

The pre-post testing method of testing team skills generally suggests that an individual will respond more accurately in the post-test because they "overestimate" their initial skills, and that following the skills assessment, the user will have met a "goal of greater understanding". Taking part in the skills assessment gives them the chance to re-assess their skills more accurately and effectively [1]. Additionally, according to [26] people have an inclination to be "blissfully unaware of their incompetence"; and this is because people establish these views about themselves on some preconception they have about the skill, not on their "objective performance".

Insights, a company focused on learning and development, have an assessment tool "Navigator Team Effectiveness" which is used to analyse the strengths and challenges of a team using a forty-two-statement questionnaire. Insights team assessment tool looks at sixteen key areas which are "critical success factors for high-performing teams" [27] which are decision making, objectives, accountability, vision, approach to task, problem solving, roles and responsibilities, communication, creativity, involvement, review of processes, meetings, trust, atmosphere, feelings, and valuing differences.

The Belbin team-role assessment is, at present, a paper-based and online assessment tool for team profiling. Belbin team roles looks at the facilitation of a team's success by portraying an individual's pattern of behaviour as it relates to their role within a team, and how they behave as part of a team. The Belbin team role assessment tool allows the team to utilize skills a team member already possesses, without having to acquire new or adapt skills to match up with current demands.

According to the Belbin team-role theory there are nine possible team roles an individual can play: plant, resource investigator, coordinator, shaper, monitor evaluator, teamworker, implementer, completer finisher and specialist. Each of these roles has their own strengths and weaknesses, and each is critical in demonstrating the unique dynamics of any given team. [27] looked at the Belbin team roles profiling assessment tool and argued that "in terms of project teams, there is no doubt that there are key group-person factors which will ensure the success of the project." It was also shown that although people do have a tendency to take on more than one role, this is typically because there is a need for them to perform more than one task. However, as [28] said, "the research has supported Belbin in his assumption that the roles are separate entities". Project leaders are going to be more likely to take on people who are willing to adopt specific roles, which is probably their preferred role, or the one which they have a "natural ability" for. In a more recent study, [29] found that by utilizing the Belbin team roles assessment tool on "a small team developing eLearning content can make a significant difference to its functioning and smooth running and…should be considered in the early planning stages of such a project." Additionally, [30] looked at comparing an individual team member's behaviour with their Belbin team role. For the research study, [30] took the executive team of the Students' Association at Stirling University and made a comparison between the individuals' behaviour within a genuine meeting room environment, and found their behaviours could be explained by their Belbin team role profiles, making predictions of their behaviours possible, with two-thirds accuracy, and "adequate validity".

Infiniteams is a customizable, online multi-player learning environment, which is designed to allow teams to participate in a series of scenarios to improve the working dynamics between each of the team members. Providing support to enable real-

time facilitation and assessment, Infiniteams has the capabilities to improve team dynamics and performance. For this study, the users were stranded on an Island, in teams of four, and working together in that team they had to collaborate with each other successfully to complete the game and get safely off the island. Infiniteams has the benefit of allowing remote collaboration, being cost-effective and customizable with the potential to implement a top-class team assessment engine.

The aim of the current research was to attempt to represent in-game scenarios, from TPLD's Inifiniteams Island game, within the Belbin questionnaire, however due to technical and experiential difficulties the research project was re-worked at the last minute. The final aim was to collect as much potential data as possible on the basis of the game, feedback, observation of success and the participants' user experience with games. Questionnaires data on group skills was also collected based on five key areas of team skills: team work, communication, leadership, problem solving and decision-making. The questionnaires were administered before and after playing the game and it was predicted that post-questionnaire responses would be more "negative" or lower than the pre-test scores, suggesting a more accurate self-evaluation of their team skills due to the over estimation people normally have of their team skills. This paper also set out to examine things that can go wrong when attempting to carry out research on team-based learning through games, and how these can be successfully overcome using the appropriate technology and methodologies.

II. METHODOLOGY

A. Design

The study adopted a within -subjects design where the dependent variables were the questionnaire scores before and after playing the games.

B. Participants

Two hundred and four psychology students from the University of the West of Scotland (UWS) (Paisley campus), from first year and fourth year classes were recruited for the study. There were no age or gender restrictions. 77% were female and 23% were male. Students were primarily of British nationality. All participants were over eighteen years of age. Using seminar slot times, students were asked to participate in a short study to pilot a new game being developed by TPLD.

C. Materials and requirements

To successfully carry out the study, participants required a computer system of their own, (which were supplied in a University computer lab) with Internet access. Additionally, another system was required by the researchers in order to set up a local server, where all graphics and game items were stored, from which to run the game. Minimum system specifications for running the Infiniteams package are: a 700MHz Processor, 256MB's of Memory/RAM, Windows 2000 or greater, DirectX 8.1 or better, headphones (or speakers), 200Mb Local Disk Space, an Internet connection, and a mouse.

The team skills questionnaire had to be uploaded onto SurveyMonkey.com, and links to this site provided to the students. The questionnaire examined five key areas of team skills: team work, communication, leadership, problem solving and decision-making. Team work included questions such as "I will be an effective contributor to the team" (pre-questionnaire) and "my contributions to the team were effective" (post-questionnaire), communication had questions like "I will communicate effectively", leadership looked more at "I will encourage the team if we get into difficulty", problem solving asked questions like "I am very adept at solving problems" and decision-making focused on questions such as "I plan to make a lot of decisions in the team". Each question, as demonstrated in the teamwork example had a post-questionnaire equivalent.

D. Procedure

The researchers used the Paisley UWS campus to carry out a short study based on their Paging Engine Infiniteams game developed from their IP2 platform, over multiple sessions, with the help of a lecturer at the University. A University lab room was used to set all the students up together, enabling them to work in teams. One system, within the lab, was utilized to set up a local server, which the game ran from. This was setup by the TPLD research staff, and one of these members acted as a facilitator for the students in the event of any technical problems arising, or to answer any questions the students may have had. Each student was then required to use there University login details to login to one of the systems, before being asked to assign themselves into teams of four, where each team member was aware of who else was in their group. Once everyone was logged in, they were asked to visit the Eduteams site (http://www.eduteams.com), where there were links to the questionnaires, and were asked to visit the link to the first, pre-game, questionnaire which re-directed them to SurveyMonkey.com where the questionnaire was hosted. After all students had completed the survey, which also asked them about their demographics – age etc - they were then directed back to the Eduteams site, where there was a link to the IP2 platform game. Students were then required to input their names, and a corresponding team code given to them by the facilitators (all team members of the same team input the same code, and all team codes were different). The students were then asked to work in their teams of four to work their way through the game and solve any challenges they faced. The game was incomplete at this stage – but this was pointed out to all the players before they started. Researchers asked the students to use the in-game chat system to allow communications to be logged. The teams had, on average thirty minutes to complete their way through the game. Once each team had reach the end of the prototype game, they were asked to return to the Eduteams site to follow the final link to the post-game questionnaire, where they were given the second part to the questionnaire assessment, and also asked for some feedback on the game they had just played, as well as some information on their gaming habits. Students who had participated were then thanked for their participation in the study.

III. RESULTS

Of the respondents, it was found that approximately 23% were male and 77% female, with the mean age of participants being 23.

All the responses from the participants were coded, with strongly disagree scored as one, up to strongly agree being scored as five, and input into SPSS. The mean scores on the team questionnaire were computed and it was found that the pre-questionnaire mean was greater (M=3.64, SD=0.39, N=177) than the post-questionnaire mean (M=3.45, SD=0.59), indicating that the responses to pre-questionnaire were more "positive". A paired samples t-test of the means for the pre-questionnaire and the means for the post-questionnaire was significant [t (176) = 3.34, p = 0.001; p<0.05], demonstrating that respondents gave significantly more favourable evaluations of theory team skills prior to participation in the team game.

Figure 2 shows the mean frequencies for the different questionnaire response categories (Strongly Agree, Agree, Neither, Disagree and Strongly Disagree) for all 20 questions, before and after, participating in the game. The graph shows that participants were more likely to agree that their group skills were good before participating in the game and more likely to disagree or strongly disagree after participating in the game.

To examine whether differences in evaluation before and after taking part in the games were evident for all team skills, the sub-scales of the questionnaires were analysed. Looking at teamwork, it was found the mean for the pre-questionnaire was greater (M=3.99, SD=0.43) than the post-questionnaire (M=3.78, SD=0.67). The mean difference between the two was 0.216 (SD=0.794), and showed a significant difference [t (176) = 3.2, p = 0.000; p > 0.001], with a 95% confidence interval (0.98, 0.34) indicating that the students' evaluations of their teamwork skills fell following game play.

The post-questionnaire mean for communication skills, (M=3.74, SD=0.68), was higher than that in the pre-questionnaire (3.72, SD=0.51) but a paired sample t-test showed that this difference was not significant [t (176) = -0.17, p = 0.87; p<0.05] and a 95% confidence interval (-0.14, 0.12).

Leadership skills had a greater pre-response mean (M=3.63, SD=0.5) than post-response (M=3.46, SD=0.68). The

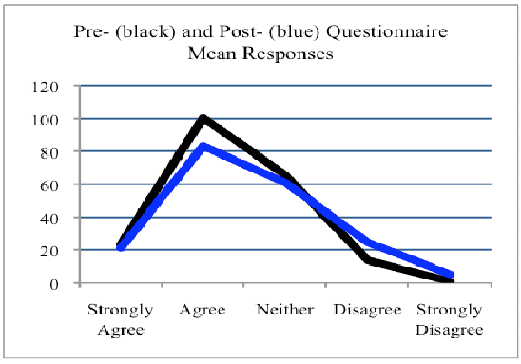

Figure 2. : A graphical representation of the mean frequency of different response categories pre- and post-game play

difference was significant [t (176) = 2.63, p = 0.009; p<0.05] with a 95% confidence interval about mean changes to questionnaire responses (0.43, 0.30). The results suggest that following the game, participants rated their leadership scores lower on the Likert scale.

Looking at problem solving skills, it was found that the evaluation scores before the game were, again, greater (M=3.55, SD=0.48) than the responses after playing the game (M=3.22, SD=0.66) and this difference was significant [t (176) = 5.036, p = 0.000; p < 0.001] with a 95% confidence interval (0.20, 0.45).

Finally, decision-making skills were analysed. The mean for the pre-questionnaire responses was 3.36, SD=0.48, and the mean post-questionnaire was 3.23, SD=0.7, showing the pre-questionnaire responses were greater. The t-test was significant [t (176) = 2.05, p = 0.04; p < 0.05] with a 95% confidence interval (0.00, 0.27).

The questionnaires also showed that 88.1% of the players enjoyed the game (Strongly Agreed and Agreed responses), 11.3% were indifferent (Neither responses), and 10.6% did not enjoy it (Disagree and Strongly Disagree responses). It was also demonstrated that 40.6% found the character easy to control, 23.1% were indifferent, and 36.3% did not think the character was easily controlled. Looking at the users understanding of the aim of the game, 36.9% understood what they were to do, 28.8% were indifferent, and 34.4% did not know. Finally, when asked if games could be used in learning 74.4% agreed, 17.5 were indifferent and 8.1% said they did not think games could be used in learning.

Feedback was also received regarding ideas for improving the game and any comments about the game, which ranged from "design a new one" and "Have dangerous snakes in it!" to more related comments such as "Make objectives of tasks clearer", "By easier control of characters" and "didn't see the point in typing a lot faster to just speak to one another, the instructions weren't clear enough and the movement of characters were difficult" – which are more constructive areas of feedback.

IV. DISCUSSION

A major problem which occurred before data collection began, but after game development had started, was that the Belbin assessment tool and the Infiniteams game were incompatible – the game took on a non-linear, open-ended format, and the Belbin assessment tool was closed-ended resulting in an incompatible match – and to overcome this, the game would have had to have been modified to remove the communications system and to restrain the users' control and movement of their character – allowing them only to choose from set choices and taking away their choice of free-flowing ideas. This resulted in a back-up plan being used: hypothesizing that the game would cause the users to examine and evaluate their perceived team working skills, more accurately which would cause lower responses on the Likert scale of their perceived abilities – although this does not automatically mean that the game does not encourage team work, but the opposite; the game encourages teamwork, and

identifies, to players, areas that they may wish to improve upon.

The results primarily support the hypothesis. Overall the mean score on the pre-questionnaire was significantly higher than that on the post-questionnaire showing that players were more confident about their group skills before participating in the game. This supports work by [1] that people will overestimate their skills initially, but also demonstrates that the infiniteams teamwork game encourages teamwork, and identifies to players, areas that they may wish to improve upon – giving them a more realistic self-evaluation of their skills – and does this not only by making the player use these skills, but by allowing them to have a greater understanding of the skill and how it can be used.

What these results demonstrate is the ability for serious games to provide an immersive environment whereby users collaborate and work together to achieve a common goal, which allows them to not only enhance their team skills, but also accurately evaluate them. This means that not only is the hypothesis supported but the underlying purpose of the game – to enable users to work in teams, analyse their team skills and to work on these skills – is also supported. If this was not the case then there would have been no significant results, as the players would not have seen any of the skills discussed, in the questionnaire, in the game play.

An analysis of the sub-scales of the questionnaires, found that there was this significant difference of lower responses post-game for Teamwork, Leadership, Problem-solving and decision-making skills, but not for Communication skills. This suggests that participants were confident of their communication skills before the game began, and their self-evaluations were reliable based on their in-game evaluation of this skill. This could be because they used these skills everyday when in conversation, writing and listening. However, looking at the other skills – skills not always used, or even considered at the time - were not as accurately assessed pre-game. The game, and the pre-questionnaire, caused the player to think about these skills and how they used them, in collaboration with others, and assess themselves on these skills more than they would in everyday life. These are vital skills needed to work as part of a team and it is essential that people understand these skills, and are able to evaluate themselves, and others, in these skills to work effectively as a team. The results do not necessarily demonstrate a decrease in these skills, but instead demonstrates that serious games can, initially, be used as an evaluation tool, enabling players to understand and assess their skills through game play, before using the games as a tool to improve these skills, supporting the idea proposed by [26].

Additional feedback from the users also suggested that they majority did enjoy playing the game, and that they generally thought games could be used in learning. There were mixed views however on the game play aspects – such as how easy the character was to control and their understanding of what the aim of the game was. This however, was also picked up in the feedback for game improvements with regards to game instructions and character movements, with comments like "Instructions on how to play the game", and "Better knowledge on what was to be achieved in task". A further, significant aspect that affected the game was that due to the problems experienced initially the game was not completed for the sessions – this was, however, pointed out to the users – but did still influence their feedback: resulting in a lot of comments along the lines of "get rid of the bugs" and "Make it work properly", which shifted the actual focus of generating ideas to improve the section of game play they had experienced.

Working with academic staff from UWS was a success for this project because it enabled the researchers (primarily from a game development background) access to a large sample of a student population, with two hundred students taking part. The study was, however, limited to using psychology students (as this was the field of the associated academia). This also gave the researchers the advantage of having the use of an academic whose field of expertise was psychology and had some knowledge of team role assessment tools, which enabled them to switch to a back-up plan using a different set of questionnaires. Within the research team, there was a learning psychologist, who was able to work closely in collaboration with the academic staff on the team assessment component, however, difficulties were ran into when the learning psychologist left the project a short time before the launch of the study. Not only would this have had an impact on the choice team assessment tool, but also on the overall study and the analysis of the results.

Adopting an Agile software development approach, for the first time, to game development, the team of researchers and game developers wanted to take on an approach which constantly reviewed what they had done, modify areas when required, work as a team, and achieve goals that benefitted the team and the end-user – using a "waterfall" process, much like that of EMERGO. This methodology, however, was not as highly successful for the team as it had been for [18], but the researchers put this down to a lack of subject knowledge from the developer team – the bulk of the work was largely put on one programmer, who did not have the same expert knowledge on the area of team assessment as the team psychologist or academic affiliate. Problems should be expected when entering a new field or trying a new methodology, and are part of the learning curve associated with anything that is new to an individual, and this method has since been successfully implemented into more recent projects, allowing the development team to work as a team, and getting the best possible results through the cycle process, constantly reviewing what work has been done, what needs to be done, implementing changes as and when they are deemed to be required, and delivering the software to the client on a regular basis is vital to making Agile work as a development methodology.

With regards to technology, the Infiniteams game has used their IP2 (Infiniteams Platform) to cover the multiplayer aspects of the game as well as math, graphics, path finding, rendering and animation utilities, resource management, GUI, collision detection and object avoidance, multiplayer framework, the audio engine instant messaging chat systems, dialogue engine and the scene management/paging system, which are all components of the IP2 library. Having all these components within the library enables rapid development from

concept to prototype, to final product, and reduces complexities on a practical level. What was found here also supports the findings [18] had with their EMERGO engine, with respects to its functionality and resources provided for the programmers. The IP2 engine provides the company with plenty of benefits, reducing development costs (by up to 70%), sales costs (by over 80%) reducing the development time (by over half) and reducing the number of human resources required (by over a quarter), compared to when no engine is used.

There were several extraneous variables that affected the study. The most significant of these was time limits – the development team had already began developing the game, in anticipation of the automatic Belbin assessment engine – but when it was discovered that this assessment would not fit with the current game system, the team (researchers included) had to return to the drawing board to generate another form of assessment. This process proved to be a lengthy process, and led to the game development team not having time to complete the game. A further problem that affected the development of the game was the number of game programmers working on the project – one. At first this was not viewed as a problem, until complications arose and time constraints were increased. The other problem with only have one game programmer was the implementation of AGILE software development – a team, as discovered by the researchers, is required to make it a success. Another major extraneous variable was one that affected the results analysis. Using SurveyMonkey.com to collect the responses to the team questionnaires and additional feedback seemed like a good idea initially, but when the data was downloaded and input into SPSS, it was found that some responses given did not match up with the user that had input the data (as discovered through a discussion with one participant following the event) – this meant that data could not be analysed for each individual, or even for the team differences (between the pre- and post-responses) and the only reliable results that could be analysed were over a group level. The researchers have since decided it would be more beneficial to develop their own questionnaire website and database – not only to rule out the data becoming mixed-up, but also to avoid the confusion of jumping between sites between questionnaires and game play – hosting them on their own homepage.

The study, although initial difficulties were experienced, has led to a greater success in the long term. The development of the Paging Engine was one of these successes. The Paging Engine meant that more focus could be applied to the graphical user interface (GUI) and in-game graphics. Also, following the key importance of assessment to this study, the researchers have since decided it would be more viable to create a more generic engine, which would run in parallel with the game. This would allow changes to be carried out externally, and would allow the engine to be matched to any (within reason) game – and this is something they are currently researching. Finally, another major improvement the researchers have been able to make since the running of the study was to make vast improvements to the event management, primarily the teacher-side user interface. The end products are designed for educational and corporate institutes, and the client management needed to run an event should be made as simplistic as possible. Running the study, the facilitators themselves ran into some difficulties with the game management system, thus demonstrating that modifications were certainly needed, and have since been implemented. Changes made were to the facilitator/management interface, minimizing the instruction needed, and the amount of effort needed to set up an event. The current study has also given insights for running future studies, namely providing better introductions to the events, as some misunderstanding was found in the current study, where some players believed the "team assessment game" referred to actual assessment within a University context, which caused some confusion for participants.

This study on the use of TPLD's game as a leaning tool, and the work that has been carried out on team assessment within serious games has led to some serious interest from big-name corporate companies, which has lead to discussions regarding further studies on the subject of team assessment gaming. The researchers also plan to finish the development of the generic team assessment engine, adapting it to be used in parallel with other games, and this would be then examined and evaluated for success in a future study.

V. CONCLUSION

The results support the hypothesis that post-questionnaire responses will be evaluated higher on the Likert scale (towards disagreeing with more statements). This was proposed to be due numerous to an over estimation of the players skills pre-game, and a more accurate self-evaluation following the use of these skills. The game play led to a greater understanding of these skills, with an overall greater awareness of their abilities. Additionally, the study also uncovered problems with the teacher-facilitator interface, the use of SurveyMonkey.com for data collection, the difficulties of Agile development using a team of under three programmers and the implications of attempting to utilize team-role assessments into a game already in development. However, lessons were learned from this, and improvements have since been implemented, showing initial success.

The results suggest that serious games, used as learning tools in education, could benefit greatly from a generic team assessment system, and this is a potential way forward for the serious games industry, specifically involving team-based games. Additionally, the research has demonstrated the need for easier facilitation, clear briefing before each session and the importance of in depth planning to cover as many potential problems, especially within the development stage, as well as the benefits of a cross-disciplinary team and the use of Agile software development, and the use of the IP^2 engine.

Potential avenues for future research are being considered, with possibilities including the development and testing of the generic team assessment project, evaluating the new facilitation interface for setting up events, longitudinal studies into the effectiveness of serious games as a skills-improvement and learning tool and further investigation into the current project looking at potential real-time coaching for serious games.

REFERENCES

[1] F. R. Rohs, "Response shift bias: A problem in evaluating leadership developmentwith self-report pretest-posttest measures" Journal of Agricultural Education, vol. 40, issue. 4, pp. 28-37. 1999

[2] G. Susi, T. Johannesson, M. and Backlund, P. "Serious games – An overview" Technical report: HIS-IKI-TR-07-001, University of Skövde 2007

[3] M. Zyda, "From Visual Simulation to Virtual Reality to Games" IEEE Computer, pp.25-32 2005

[4] K. Corti, "Games-based learning; a serious business application" PIXELearning Limited. www.pixelearning.com 2006

[5] B. Sawyer, "Ten myths about serious games" http://www.escapistmagazine.com 2007

[6] B. Sawyer, "The state of serious games" http://www.gamasutra.com 2005

[7] S. Egenfeldt-Nielsen, "The basic learning approach behind serious games" www.seriousgames.dk/ 2005

[8] M. Slocombe, "Casual games market growing by 20 per cent a year" http://digital-lifestyles.info October 2007

[9] A. Hutson, "Totally serious about games!" http://www.seriousgamesinstitute.co.uk 2008

[10] C. Ondrejka, "Finding common ground in new worlds" Games and Culture, Vol. 1 Issue. 1 pp.111-115. 2006

[11] S. Egenfeldt-Nielsen, "Can education and psychology join forces. The clash of benign and malign learning from computer games" http://www.nordicom.gu.se 2006

[12] S. A. Barab, K. E. Hay, M. G. Barnett, & K. Squire, (2001). Constructing virtual worlds: Tracing the historical development of learner practices/understandings. Cognition and Instruction, Vol. 19 Issue. 1, pp.47–94. 2001

[13] J. D. Bransford, A.L. Brown, & R. R. Cocking (Eds). (2002). "How people learn: Brain, mind, experience, and school." Washington, DC: National Academy Press.

[14] A. Nareyek. "Game AI Is Dead. Long Live Game AI!" IEEE. Vol. 22, No. 1 2007

[15] E. Aarseth. "Playing Research: Methodological approaches to game analysis" In W. Bo Kampmann (Ed.), Game Approaches / Spil-veje. Papers from spilforskning.dk Conference August 28-29, 2003.

[16] E. Van Damme, "Game Theory: The Next Stage, " in L. A. Gerard-Varet, Alan P. Ki rman, and M. Ruggi ero (Eds.), Economics be yond the Millennium, Oxford University Press, 1999, pp. 184-214.

[17] G. A. Gunter, R. F. Kenny, and E. H. Vick, "A case for a formal design paradigm for serious games" The Journal of the International Digital Media and Arts Association, vol. 3, pp93-105 2006

[18] R. Nadolski, H. Hummel, H. van den Brink, R. Hoefakker, A. Slootmaker, H. Kurvers, and J. Storm, "EMERGO: methodology and toolkit for efficient development of serious games in higher education" http://dspace.ou.nl 2007

[19] W. Westera, R.J. Nadolski, H.G.K. Hummel, and I.G.J.H. Wopereis "Serious games for higher education: a framework for reducing design complexity" Journal of Computer Assisted Learning. Vol. 24. Issue. 5. pp420-432 2008

[20] P. Abrahamsson, O. Salo, J. Ronkainen, and J. Warsta, "AGILE software development methods" Review and Analysis. ESPOO 2002. VTT Publications 478, 2002

[21] J. Highsmith, "What is AGILE software development?" The Journal of Defense Software Engineering, pp 4-9 2002

[22] H. Kniberg, *Scrum and XP from the Trenches*. Stockholm: Crisp, 2007.

[23] T. Marsh, K. Yang, and C. Shahabi, C. "Game development for experience through staying there" In sandbox '06: Proceedings of the 2006 ACM SIGGRAPH symposium on Videogames, pp. 83-89, New York, NY, USA. ACM.

[24] S. Chen, and D. Michael, "Proof of learning: Assessment in serious games" Gamasutra October, 2005 http://www.gamasutra.com/

[25] J. Bull, M. Danson, G. Conole, H. Davis, S. White, and N. Sclater, "Rethinking assessment through learning technologies" Proceedings of Winds of Change in the Sea of Learning: Charting the Course of Digital Learning (ASCILITE 2002) UNITEC, Institute of Technology, Auckland, New Zealand. 2002

[26] D. Dunning, K. Johnson, J. Ehrlinger, and J. Kruger. "Why people fail to recognize their own incompetence". Current Directions in Psychological Science, Vol. 12, Issue. 3; pp.83-87 2003

[27] Insights Navigator Team Effectiveness. http://www.insights.com

[28] J. Sommerville, and S. Dalziel, "Project teambuilding – the applicability of Belbin's team-role self-perception inventory", International Journal of Project Management, Vol. 16, Issue. 3; pp. 165-171 1998

[29] A. Gibson, and T. Nesbit, "Belbin team roles, organisational patterns and eLearning: A case study" http://site.tekotago.ac.nz 2006

[30] T. K. Johansen, "Predicting a team's behaviour by using Belbin's Team Role Self Perception Inventory" Unpublished. 2003