



Evaluating coursework in computer games degrees: Students and assessors as virtual characters

Barry Ip, Martin Capey, Andrew Baker and John Carroll
Swansea Metropolitan University

This paper explores a relatively new area in the design and development of assessment procedures for the evaluation of coursework and student performance on computer and video games degrees. Emphasis is placed on an assessment which involves the development of and interaction in a virtual world, where lecturers and students are represented as virtual characters. The establishment of a new assessment framework for games degrees is outlined, followed by the findings from in depth interviews with students and lecturers concerning its implementation. The results provide a self reflective but critical insight on how existing assessment techniques may be adapted for the growing demands and popularity of games courses, as well as issues surrounding the appraisal of undergraduate work within virtual environments. Recommendations are provided on how similar assessments may be conducted in the future.

1. Introduction

Games designed for use in education are commonly reported as being powerful motivators (De Aguilera & Mendiz, 2003; Amory, Naicker, Vincent & Adams, 1999; Cagiltay, 2007; Malouf, 1987; Rai, Wong & Cole, 2006), helping to facilitate the acquisition of knowledge on specific subjects as well as essential practical skills such as problem solving, reasoning, and creative thinking (De Aguilera & Mendiz, 2003; Amory et al, 1999; Pillay, Brownlee & Wilss, 1999). The aim of this paper is to extend on the existing body of work by detailing a relatively new and unique form of game and student evaluation in higher education, where a part of the assessment takes place within a gaming environment in which both students and examiners (in this case, the module lecturers) are represented and interact as virtual characters. A special consideration is placed on the development of a new assessment framework which provides a more holistic evaluation of student work in the emerging field of computer games design and production.

Numerous approaches exist for appraising the effectiveness of educational games, as extensively documented by Inglis, Ling and Joosten (2002, 165-88), Laurillard (2002, 86-89), Phillips (1997, 127-146), and Prensky (2001, 379-386). In the context of this study, three principal types of assessment hold special relevance in relation to interactive games: 1) formative techniques, such as the measurement of a game's effectiveness with respect to appropriateness or relevancy of content, ease of use, and the method of implementation within the context of the core curriculum (see, for example, Cheung & Siu, 2002; Feinstein, Mann & Corsun, 2002); McFarlane, Sparrowhawk & Heald, 2002); 2) summative evaluation, especially that of gauging any significant gains in student

performance, i.e. whether or not the use of such games leads to improvements in terms of academic achievement (Hong & Liu, 2003; Oyen & Bebko, 1996; Rosas et al, 2003), and 3) the use of lifelike virtual characters and their potential to be used within assessment frameworks. Highly detailed, animated 3D characters, often employed as teaching assistants or guides within educational games (commonly referred to as pedagogical agents), are gradually becoming more prominent in the light of advances in modern technology.

As has been recently evidenced by the expansion of research in this area, findings indicate agents can contribute towards a heightened development of students' cognitive skills, and, due to the highly believable or anthropomorphic nature of well-designed virtual agents, learners may also become emotionally attached to the characters being controlled or interacted with, thus further enhancing learning motivation (Clark & Choi, 2005; Gulz & Haake, 2006; Kim, Baylor & Shen, 2007; Moreno, 2005; Moundridou & Virvou, 2002). In addition, Moreno (2005) points out how virtual agents may, in future, play a vital role in facilitating the demonstration or simulation of technical procedures in emerging fields. Thus, one of the aims of this research is to ascertain whether the use of virtual agents offers any notable advantage during the assessment of student work produced on computer games degrees.

Useful though the above evaluation techniques are, the idiosyncrasies relating to the assessment of students' abilities on computer games courses (typically, the design, development, and production of interactive games) give rise to a unique set of challenges. The increase over recent years in the number of games development courses in UK and overseas universities – see EDGE (2006) and EDGE (2007) – reflects not only the growing popularity of computer games as an academic subject, but the games industry's fervent need for graduates with practical game design and production skills (*Develop*, 2007; Hasson, 2007; Jeffery, 2007; Sanches, 2007; Sperry, 2006). Consequently, one fundamental distinction between assessments required for games courses and the techniques described above is that the latter are designed and used principally for material (games or other forms of interactive software) intended for educational instruction or delivery, whereas the former need to be geared towards the development and production of commercial games which may not contain any educational application beyond those required for the assessment of learning outcomes. So, for example, a student may produce a shooting game to demonstrate his or her ability to create 3D models, textures, level designs, and the like, necessary for the fulfilment of specific learning criteria on a games course, yet the game is likely to contain little or no content for the purpose of educational instruction. Consequently, there is, at present, a mismatch between established methods of game evaluation devised principally for educational material, and those specifically for the assessment of game design and production.

Notable pedagogical differences have also been outlined by Rieber (2005) between games and traditional multimedia or interactive material. It is stated that educational multimedia centres predominantly on instruction or explanation of some predetermined content, whereas games focus primarily on providing a particular experience or sensation of play (Rieber, 2005). Hence, in relation to a formal analysis of games, despite growing work in the field (see, for example Consalvo & Dutton, 2006; Ip & Jacobs, 2004; Malliet, 2006; Sweetser & Wyeth, 2005), a great debate remains on how games may be analysed scientifically in terms of experience, enjoyment, or how much fun they are to play. An important issue has also been alluded to by Laurillard

(2002, p. 206) in the context of how the evolution of learning technologies alter the way in which assessments should be carried out, and the key responsibility of teachers to judge what constitutes 'good' performance in relatively new subjects. Thus, taking into consideration of the differences between educational games and commercial games, between games of experience and games of instruction, and the uncertainty of how best to gauge student performance in emerging topics, any attempt to apply established assessment techniques for the evaluation of student performance in the field of games design without careful adjustment for relevancy and structure would be misguided, if not largely inaccurate.

The case study presented in the remainder of this paper addresses some of these issues by examining how an extensive modification of existing assessment techniques was made for a new computer games degree. The findings also provide a unique insight into how use of virtual agents – here, the representation of students and assessors as virtual characters – combined with established assessment techniques, can enable lecturers to obtain a more appropriate and accurate judgement of student performance in the growing field of games design. From here, recommendations are provided on how future modifications may further enhance the assessment process.

2. Game components and virtual characters

The game examined as the central part of this research was developed by a group of three undergraduate students as a major project over a 12-month period between June 2006 and June 2007. The major project module represents two-thirds of the students' entire final year coursework as part of a BA degree in Creative Computer Games Design (CCGD) at the School of Digital Media in Swansea Metropolitan University, in the UK. The group of three were among the first graduate cohort consisting of 10 students in total, with the course being initially established in 2004 as one of the first dedicated games development courses in the UK. In contrast to games courses emanating from computer science disciplines, CCGD approaches game design from an art and design perspective, focusing (though not exclusively) on artistic, ludological, production, and narrative components of game creation. The course contains core areas of study recommended by the curriculum framework outlined by IGDA (2003). All students in the cohort received weekly tutorial support for the major project from four key members of the lecturing team on the games course, and were directed towards practical aspects of game development and production including project planning, game conceptualisation, 2D and 3D animation, game engines, and level design, to reflect the syllabus as well as specific demands of the industry. The only stipulation in terms of the production of the game was that it had to be interactive.

The final game produced by the three students, named *Ertha*, was produced using Epic's *Unreal Tournament Engine 2004*. The general idea behind *Ertha* is to demonstrate how future commercial, massively multiplayer online games may be developed, with particular attention on the use of user generated content. The salient aspect of *Ertha* in the context of the assessment was the representation of all three students and four assessors (the course lecturers) as virtual characters – these are shown in Figures 1 and 2 respectively. As can be seen, the students modelled themselves as tour guides, and would provide directions and game instructions to the lecturers during the assessment (discussed further below). The use of virtual agents was originally proposed by the students prior to Stage 4 of the assessment (specified below) as they felt it would enable the lecturers to gain a much clearer understanding of the finished game. This

idea was unique to this group of three students, and whilst other projects produced by the remaining student cohort were all interactive and functional, *Ertha* was the only game in which students and lecturers were incorporated in such a way.



Figure 1: In-game character models of students



Figure 2: In-game character models of course lecturers/assessors

The final version of the game submitted for the assessment contains a tutorial level, level hub, five core levels, and three sub-levels, structured in the layout shown in Figure 3. As can be seen in the diagram, the game begins with a tutorial section which explains the basic controls from where players are directed to the central hub. Here, players are able to access any one of the five core levels and teleport between them. Three further sub-levels may be entered via the City. Figures 4 to 6 below provide examples of the visual style contained in *Ertha*.

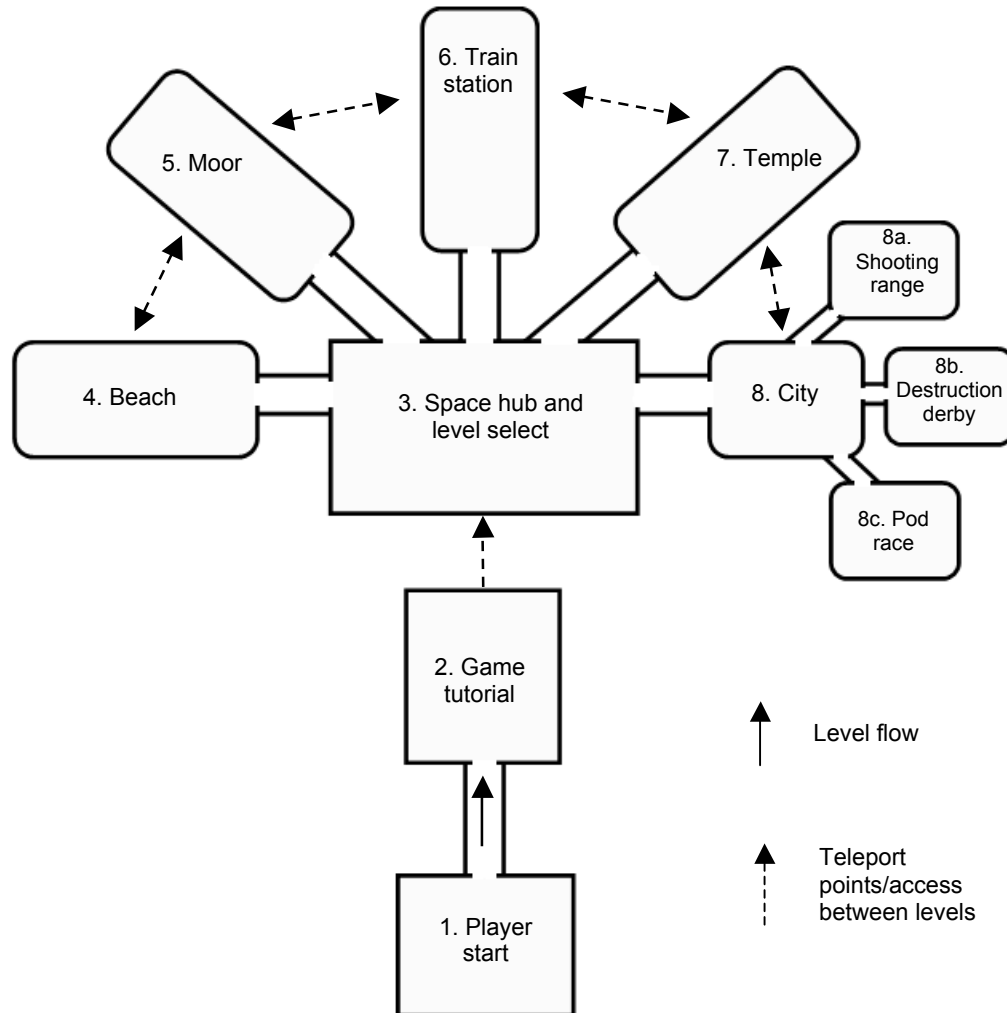


Figure 3: Basic level map for *Ertha*

3. The assessment

As discussed above, existing game evaluation methods required careful adjustment such that the appraisal of learning outcomes, along with specific aspects of games development, could be suitably incorporated into the assessment procedure. This



Figure 4: Screenshot of space hub and level select

particular challenge was most prevalent for the major project, which being run for the entire duration of the final academic year, represents a sizeable investment of students' time and effort, requiring them to evidence a broad range of practical and theoretical skills necessary for successful completion. Through extensive consultation with academic subject specialists and industry experts during the degree validation procedure, the decision was made to implement a tiered assessment structure, consisting of four stages, to take into account the wide spectrum of work produced by the students throughout the year, as well as the relevant learning outcomes to enable accurate assessments to be made.

Established formative and summative evaluation techniques were used throughout, including a demonstration/play-test of the completed game in the final stage (see Table 1). Appropriate modifications to assessment type were made depending on the assessment stage (and the associated learning outcomes), with each stage being evenly spaced to allow students adequate time to develop their project, but considerable attention was placed on ensuring other relevant techniques of evaluation for games development are also included. For instance, weekly tutorial sessions focused on gathering feedback from students and the lecturing team about course delivery and assessment stages, as suggested by Fry, Ketteridge and Marshall (1999, 161-74), in addition to providing students with the necessary supervision on the development of their game, design evaluation, and needs analysis (Inglis, Ling & Joosten, 2002, 168-72). Due to the iterative and design-based nature of the work, students were given the opportunity to engage in weekly discussions and self- and peer-reviews of their work in order to facilitate a reflective approach towards learning, as recommended by Elwood and Klenowski (2002), Hansen (2004), and Langan et al (2005).



Figure 5: Screenshot of temple level

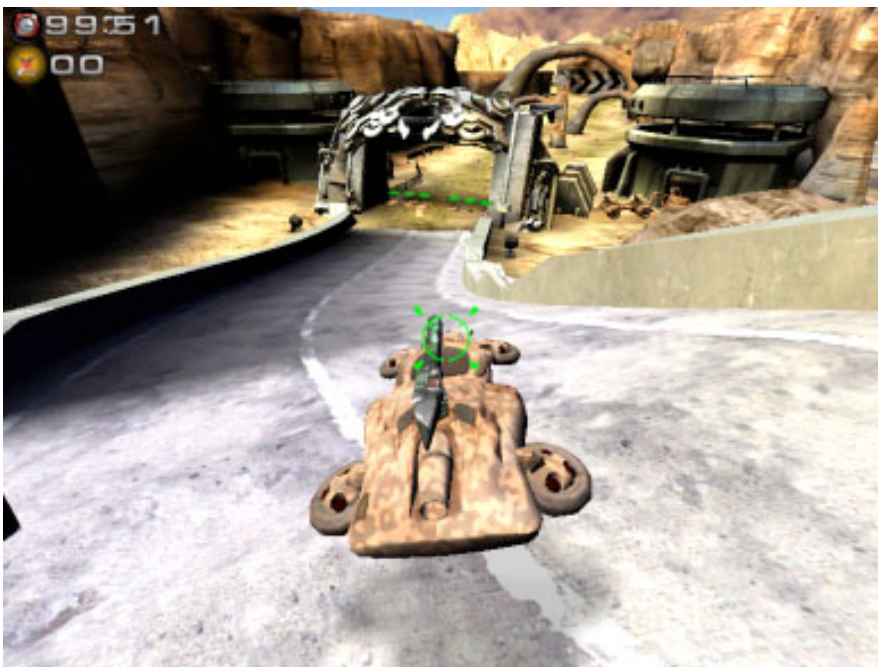


Figure 6: Screenshot of pod race level

Table 1: Overview of assessment stages

Stage (weighting shown in brackets)	Assessment format	Assessment type	Assessment time and duration
1. Project seminar (15%)	Students present a topic related to a proposed game idea.	<p><i>Formative/supervisory support:</i> Weekly supervision and tutorial sessions (around four hours per week, plus any additional support students may require) for feedback on current stage of assessment, student progress, general course-related issues, relevancy and appropriateness of proposed topic.</p> <p><i>Summative:</i> 1. Assessment criteria were designed to evaluate the students' fulfilment of specific learning outcomes – presentation skills, ability to research the topic, clarity of material, etc. Students were awarded marks according to their ability to evidence the skills necessary for the current stage of assessment. 2. Submission of all presentation material.</p>	Viva held towards end of first teaching term: approximately two months after commencement of academic year; 30 minutes per student (including 15 minutes of discussion/questions).
2. Pitch (15%)	Students compile, present, and submit a detailed project plan to obtain approval for a game proposal.	<p><i>Formative/supervisory support:</i> Weekly supervision and tutorial sessions (around four hours per week, plus any additional support students may require) for feedback on current stage of assessment, student progress, general course-related issues, guidance on appropriateness of proposed game and methods of implementation (e.g. game engines, application of game theories, timeframe).</p> <p><i>Summative:</i> 1. Assessment criteria including justification of project proposal, evidence of project planning, and conceptual skills. Students were awarded marks according to their ability to evidence the skills necessary for the current stage of assessment. 2. Submission of written report (typically, a game concept document and project plan) and all presentation material.</p>	Viva held at the end of first teaching term: approximately four months after commencement of academic year; 30 minutes per student (including 15 minutes of discussion/questions).
3. Interim (10%)	Students present all work-in-progress before final submission.	<p><i>Formative/supervisory support:</i> Weekly supervision and tutorial sessions (around four hours per week, plus any additional support students may require) for feedback on current stage of assessment, student progress, general course-related issues, additional staff and student meeting for official feedback on student progress, course delivery and structure, guidance on work-in-progress (playability and technical issues, testing, documentation).</p> <p><i>Summative:</i> 1. Assessment criteria including evidence of technical ability, transferring of concept into final product, and problem solving/design iteration. Students were awarded marks according to their ability to evidence the skills necessary for the current stage of assessment. 2. Submission of written report (typically, a detailed game design document) and presentation material.</p>	Viva held at the end of second teaching term: approximately seven months after commencement of academic year; 30 minutes per student (including 15 minutes of discussion/questions).

4. Final product (60%)	Students present and demonstrate their final product/ game.	<p><i>Formative/supervisory support:</i> Weekly supervision and tutorial sessions (around four hours per week, plus any additional support students may require) for feedback on current stage of assessment and support on final product viva.</p> <p><i>Summative:</i> 1. Assessment criteria including quality of final product, ability to manage project according to timeframes, and level of contemporary technical insight in the development of the final product. Students were awarded marks according to their ability to evidence the skills necessary for the current stage of assessment. 2. Submission of final product, written report reflecting on entire project, and presentation material.</p> <p><i>Game testing:</i> In the case of <i>Ertha</i>, the use of virtual agents where students and assessors (the four members of the lecturing team) are represented as virtual characters within the game. Students' abilities in the production of an interactive game were gauged through the gaming experience and all submitted material.</p>	Viva held in final teaching term: approximately nine months after commencement of academic year; 30 minutes per student (including 15 minutes of discussion/ questions).
------------------------	---	---	--

All four stages of assessment were conducted by the same four members of the lecturing team. It was felt that being such a new course, the full participation of the four key members of the teaching staff was essential as they each specialised in unique areas of games design and production, and therefore would enable all aspects of student work to be judged with equal emphasis. Students were required to produce and submit the pre-specified written and practical work for each stage, culminating in a viva where each student would present and/or demonstrate the relevant work (the weightings for each viva, of which are split according to the appropriate assessment criteria, are outlined in Table 1). The type of work produced by students would evolve throughout the stages to reflect the various steps in the game production pipeline: a largely conceptual and theoretical basis at Stage 1; a mixture of practical and theoretical approaches outlining a game proposal in Stage 2; mostly practical work, including a significant development of game assets and supporting documentation in Stage 3; completion of a final product in Stage 4. As detailed in Table 1, great effort was made to ensure the vivas adhered to the suggestions for best practice in oral examinations, as described by Davis and Karunathilake (2005) and Guest and Murphy (2000), including clear and pre-specified assessment procedures, the use of a series of examinations (in this case, four), consistency in questioning (determined by the assessment criteria), and the use of explicit assessment criteria.

In Stage 4, where the finished game is demonstrated, students provided lecturers with playing instructions, detailed explanations of each section of the game, and information on its construction, design iterations, and functionality. The role of the lecturers was to observe, interact with and evaluate the game, as well as to raise any questions concerning the development of its content. Once the presentation and discussion were completed, the assessors met in private to consider the performance of the students and the quality of the final product (the time permitted for this is shown in Table 2). Marks were then allocated for the various assessment criteria defined in

the course document. These procedures were applied to all 10 students in the cohort. For Stage 4, a few notable additions were made to facilitate the use of virtual agents specifically for the case of *Ertha*. Seven PCs were networked to enable the students and lecturers to partake in the game collectively. Headsets with built in microphones were used to allow students to provide verbal instructions (such as for directions on where to navigate, control layouts, and recovery procedures) to the lecturers during the game. At the start of the assessment, the three students and four lecturers commenced the game simultaneously in the *Player Start* section and progressed through each stage in the numerical order shown in Figure 3 (thus beginning from *Player Start* and finishing in the *City*) to enable all parts of the game to be examined.

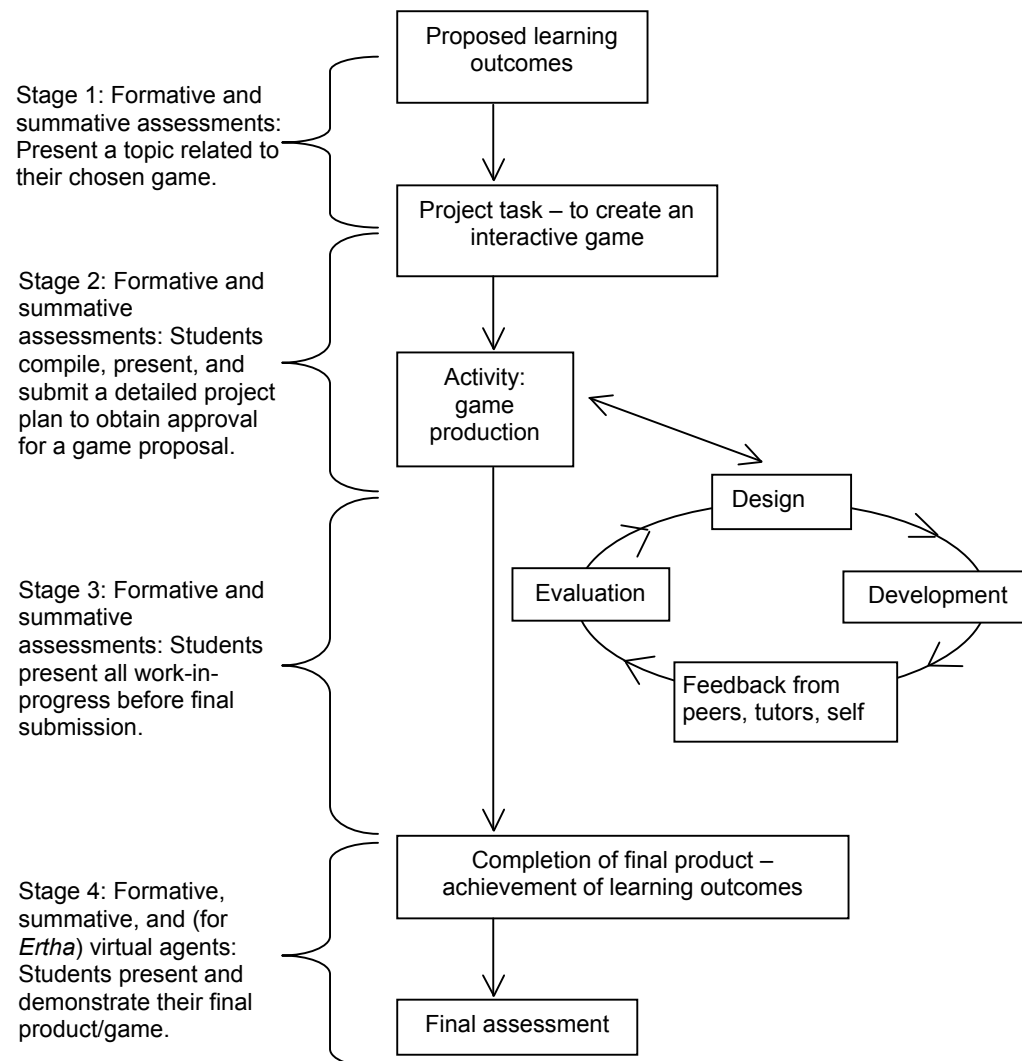


Figure 7: Overview of assessment stages.
Adapted from Inglis et al (2002: 169), Maier and Warren (2000: 133), and Phillips (1997: 38)

In summary, the entire assessment framework undertaken follows the structure conceptualised in Figure 7. It can be seen that in addition to satisfying the necessary academic requirements for modern assessment practices (Elwood & Klenowski, 2002), the process adheres to existing evaluative models proposed for digital, interactive product development cycles (Inglis et al, 2002, p. 169; Maier & Warren, 2000, p. 133; Phillips, 1997, p. 38), hence providing students with experience of work schedules likely to be encountered in games and games related industries.

The following section outlines student feedback on the general assessment structure from the entire cohort, and those specific to *Ertha*, obtained from the three students and four assessors using exploratory, depth interviews supported by a structured questionnaire. All interviews were conducted by the same interviewer from the research team and recorded with students' and lecturers' permission. Interviews were subsequently transcribed for further analysis. The research undertaken combines the approaches prescribed by the principles of self reflective action research and structured interviewing. A copy of the interview questions is given in the Appendix.

4. Feedback on the assessment

The summary of findings is split into two sections, covering attitudes relating to the general structure of the assessment (gathered from all 10 students of the cohort) and the use of virtual characters in *Ertha*.

4.1. General assessment structure

Feedback from students regarding the assessment was largely favourable, but two main aspects of concern were highlighted: the provision of adequate time for each assessment stage, and for those working in groups, the need for lecturers to gain an understanding of students' varying roles within their projects.

Whilst there was collective agreement by students that the duration of 30 minutes per student was fair and appropriate, responses from the lecturers draw attention to a plethora of key issues concerning this aspect of the assessment. First, it was felt that the current duration of assessment was made possible largely by the relatively small student numbers (10) in the 2006-2007 graduate year of the CCGD degree. All four lecturers (passionately) agreed that the viva duration and structure should be maintained, but one response offered a particularly insightful summary of the current situation and potential direction of future assessments:

There is no way we could consider assessing the work without vivas – it is essential for this sort of subject, which also includes multimedia and, to a smaller extent, 3D animation. This is due to the nature of the work that is likely to be produced. In the final year of the games degree, students will only have enough time and resources to complete a partially completed game – realistically, they will never be able to produce one completely finished piece of work. So, the vivas are essential for students to explain how the game fits a certain context, on the basis that it will be incomplete, just like how an early game idea would be pitched to a game investor in the industry who might decide to take it on and spend another two years on it. Until the time comes when the technology gets to a point where it will allow students to make a complete game in a realistic timeframe, which the lecturers can operate seamlessly with little or no explanation, the vivas will be imperative to the assessment.

Essentially, the view offered here is that vivas are essential not only for assessment purposes (in terms of length and structure), but are necessary due to the nature of current game development technology. The critical point lies in the resource intensive nature of current games development, and hence the expectation on students to produce fully completed and self explanatory games (which, generally, consist of large expansive environments) during a relatively short time period is largely unrealistic. So long as this is the case, students should be given the opportunity to explain and contextualise their work on the basis that it is likely to be only partially complete.

One the whole, it is feasible that the current duration of 30 minutes per student can be efficiently maintained for future cohorts consisting of no more than around 30 students on any given degree pathway in the School, as is presently the case for courses such as BA and BSc 3D Computer Animation, BA and BSc Multimedia, and BSc Music Technology, where identical viva durations are timetabled annually. This provision, however, will come under threat if student numbers were to substantially increase. All lecturers agreed that vivas are essential, but no clear answer could be found for the questions associated with the rise in student numbers and the inevitable impact this would place on resources. To illustrate the possible effects, an illustration is provided in Table 2 outlining the time requirements for the current scenario at the University, a forecast for future cohort sizes, and possible cohort sizes at institutions with larger student numbers.

Table 2: Current and projected time requirements for vivas

Cohort size	Present- ation duration per student (mins)	Discussion time between assessors (takes place immediately after each presentation) (mins)	Total viva time per student (mins)	Total time required for cohort (hrs)	Days required to assess cohort, based on a max. of 7 vivas per day, requiring approx. 5.25 hrs per day
10 (2006/07 grad. yr)	30	15	45	7.5	1.4
30 (expected cohort for 2007/08 grad. yr)	30	15	45	22.5	4.3
50 (possible cohort size for future yrs)	30	15	45	37.5	7.1
100	30	15	45	75	14.3
200	30	15	45	150	28.6

As can be seen in Table 2, future cohorts of 30 students (currently the case for first and second year students) require just over four working days to assess using the current viva structure. Thus, vivas for all four stages of assessment would require no more than four assessment weeks, spread across the entire academic year. Resources will become stretched, however, for cohorts of 50 or more. For example, for 100 students, the 14.3 days required for vivas may be feasible as an end of year assessment (around three weeks would be assigned specifically for this), but considerable disruption would be inflicted on the rest of the academic year if the four stage structure is retained (which would amount to a total of around 12 weeks being assigned to vivas). This not only places a considerable workload on lecturers/assessors, but would reduce the amount of teaching and supervision time that would otherwise be allocated. Interestingly, even with the 1.4 days required for this cohort, one student reported that staff appeared to be overworked throughout the year, and would have liked to see additional staffing to alleviate the workload on the four principal course lecturers.

One of the most apparent remedies for coping with significant increases in student numbers is to reduce the number of assessors in each viva (say, from four to two), and to then utilise staff across parallel assessments. Thus, providing sufficient viva resources (such as rooms and computer equipment) are available, the assessment durations shown in Table 2 may conceivably be halved. In this scenario, a cohort of 100 students will be split into two streams, each assessed by two lecturers, therefore requiring just over seven days for each stage. This approach could therefore alleviate any excessive impact of vivas on the teaching term. However, while there was a strong degree of agreement in the assignment of grades between the lecturers for 2006-07 graduate year (in cases where agreement was weak, lecturers deliberated to arrive at a consensus), future research must examine levels of inter-rater agreement in greater detail to ensure consistency in marking between the different sets of assessors before parallel assessments can be adopted.

Alternatively, the assessment may be performed independently by individual lecturers (each being an expert in a particular aspect such as design, technical, artwork, and gameplay), evaluating only learning outcomes specific to their field. While this approach may provide some additional time economies (as lecturers do not necessarily have to be scheduled together, and may assess the work outside of a fixed viva schedule), its practical application may be problematic since, unless alternative meetings are scheduled separately, students will not have the opportunity to explain the work to the lecturing team as a single entity. However, with the potential for sizeable increases in student numbers, this approach potentially offers a practical compromise for assessments. Clearly, this remains a difficult issue on which there appears to be no simple answer, and will be returned to later on in this discussion.

Positive responses were received from both lecturers and students on the formative assessment stages/weekly tutorial sessions in which students were given regular feedback on the progress of their work, as well as explanations on specific assessment criteria before each viva. These findings support the findings from research into similar assessment formats (Carless, 2002). For the lecturers, while students were not formally assessed as in the summative stages, the formative feedback sessions proved extremely valuable in gauging students' attainment (or otherwise) of learning outcomes on a weekly basis. This not only enhanced lecturers' understanding of the projects and technical aspects of students' work prior to each assessment stage, but also helped to identify areas where extra supervisory support might be required. Furthermore, regular feedback meant the lecturers gained a stronger appreciation of the varying roles (such as game designer, technical lead, 3D artist) the students played within their team, and hence of how each student had specialised on a particular facet of games development. Thus, by the time of the actual vivas, the lecturers felt they were able to judge student work with much greater accuracy. Feedback from the students was also encouraging. They commented on regular tutorial sessions being of great benefit in terms of alleviating their concerns about group work (for instance, in the event of group members not pulling their weight), the issue of specialisation and workloads within the project, and how their specific skills would be taken into account during assessments.

With regard to the Stage 4 viva, lower levels of anxiety were reported by the students. Specifically, and somewhat surprisingly, the students remarked how much they enjoyed this particular stage as compared to previous assessment stages. Notable comments made by students included:

Comment 1:

I much preferred doing the assessment with the lecturers sat at computers as we guided you through. It felt fun and informal and so I didn't feel nervous in the way that I did in the previous vivas. Standing up and doing a presentation in front of four stern faced lecturers is pretty intimidating. Sitting down and playing games with four bemused faced lecturers is fun.

Comment 2:

The final viva was for me a lot less daunting than any other viva, this was definitely due the fact we were playing games. In previous vivas there was a very obvious barrier stopping students and lecturers interacting with each other, with a different level of formality and professionalism. In the final viva, we were able to interact and be removed from the tense environment of traditional presentations. For me, this was incredibly useful. At times, it didn't feel like I was being examined, and instead we were just showing you what we had done.

Comment 3:

Once the viva was underway and the lecturers began to respond, the whole thing became far more instinctive and natural than previous vivas. Having everyone absorbed with the virtual environment allowed for a more relaxed method of communication and took a certain undesirable formality away from the viva. Hearing and watching the lecturers interact and play was incredibly gratifying.

Clearly, the interactive element of the Stage 4 viva appears to have significantly enhanced students' experience of the assessment. These results were encouraging, though unexpected, in the sense that with the exception of the interactive session, the Stage 4 viva was conducted in an identical fashion to the previous stages in terms of structure and formality, yet the students felt much more positively. On the whole, the lecturers did not feel the viva was less formal than previous assessments. As with the other three stages, students were clearly informed of details such as viva structure, time restrictions, and the allocation of marks, but it offered lecturers the unique opportunity to experience the final product. Indeed, one lecturer remarked that this particular stage actually felt more formal than some of the previous assessments as the students had appeared to make extra efforts to make a structured and professional presentation, as would be expected in industry.

All the lecturers agreed that the lower levels of anxiety felt by the students was a positive outcome, and should be fostered, as much as is appropriate, for future assessments. However, the issue of vivas becoming 'too informal' did raise some concerns for one lecturer, who felt that the assessors should have something more substantial to contribute to future vivas, rather than to merely interact with a game. The following section elaborates on this issue.

4.2. Interactivity and the use of virtual characters

This section applies specifically in the context of *Ertha*, where the virtual characters of both students and lecturers (Figures 1 and 2 respectively) were used for the assessment. The lecturers' interaction with *Ertha* during the Stage 4 viva was considered by the three students to be an imperative component of the assessment. This particular view was aptly summarised by one student:

With the predominately interactive nature of the videogame, we felt that to passively observe the final build of *Ertha*, as opposed to playing it, would undermine the product. Although a pre-rendered animation can be assessed in this manner, a

videogame needs interaction in order to comprehend what has been achieved. That was why the lecturers should be involved in the final viva.

With regard to the interaction in *Ertha*, students were concerned with their decision to present the game in a linear fashion, that is in the order shown in Figure 3. The reason for this is that *Ertha*, being designed as a massively multiplayer online game, would ideally be played without any player restriction, as real gamers would experience it. However, a non-linear, free roaming approach would likely increase the possibility that critical parts of the game are overlooked, and hence students decided to impose restrictions to ensure lecturers could experience (and therefore assess) every important element contained within the game. On reflection, the linear approach was preferred by the lecturers, who felt it to be essential not only to ensure all facets of the game would be assessed, but also to explain and contextualise the game's content and provide a direction for the interaction. Indeed, the lecturers remarked that without the explanations presented in the viva, they probably would have awarded slightly lower grades had the work been marked in isolation from a practical demonstration. Time permitting, future vivas may incorporate the use of a free roaming, introductory play session followed by a linear, guided demonstration to take advantage of both approaches.

The students felt that the use of virtual characters offered a unique way of personalising the viva experience. They remarked that the intention of using specific characters was to strengthen lecturers' physical involvement with the demonstration. However, the students also commented that the application of the characters was context specific, i.e. it may not be appropriate to use representations of lecturers and/or students if they do not fit the environment (e.g. a war simulation, horror game, or fantasy). Overall, lecturers certainly agreed with the notion that the 3D characters enhanced the degree of immersion in the game. The characters enabled the lecturers to become an integral part of the virtual world, being able to observe and manipulate it from within, as opposed to a greater degree of detachment in other game assessments where there was no physical representation. As opposed to the students' views, the lecturers felt that the most powerful aspect of the characters was that they were modelled as the students and lecturers, and thus contained direct relevance to the assessment. This particular aspect also helped the lecturers to take a more objective view during the demonstration, reminding them of their role as assessors. In addition, one lecturer drew attention to the transfer of the physical aspect of assessment from the traditional classroom to the virtual space, and its potential ramifications:

I thought the in-game characters fitted the context perfectly. As a viva/assessment experience, I've never experienced anything like it before. I think it opens up a whole new concept on how work and material such as this can be examined in the future. This sort of thing has been done for a few years on commercial games, where you can make a version of yourself to play in the game, but to actually apply it in the context of an assessment is a whole new concept. One of the things it did was to transfer a lot of the structure and formality of the assessment from a physical location (e.g. a lab or lecture room) into the game. When everyone got in the game, saw each other, and walked around, the issue of assessment was very much left in the background. It was only when the game stopped that we took a step back and thought critically about the work.

From this, the final salient point derived from the interviews relates to the development of game content and lecturer participation in future assessments. The lecturers felt that students should not be given the ultimate say in determining game

content without staff approval, an aspect in which students are currently afforded a large degree of freedom. The concern resides in the likelihood that students may portray sensitive and potentially inappropriate material, for example, that which contains excessively violent content which is often reported in the media. Therefore, the challenge for the future is to determine how best to balance students' creative freedoms and supervisory control to encourage themes beyond those which might be considered as distasteful or, at the very least, stereotypical of the industry. Indeed, these issues reflect measures adopted by the games industry itself for the regulation of game content and the establishment of age ratings (see ELSPA, 2007). In addition, one potential requirement of the assessment might be the use of more explicit project criteria to provide greater direction for the design and development of game interaction, through which a greater degree of control may be exerted. On this issue, two lecturers commented:

I think we, lecturers, should have something to contribute to the virtual environment rather than just being reactive. So, for example, we might ask students to enable us to bring things into that virtual space for the purpose of the assessment. I think there are new methods and protocols which can be established to further enhance assessments for this sort of medium.

And, on a more general note:

I agree in large part with the idea for the future where students should set lecturers some sort of ultimate objective in the game, to make us play a more central role in it. Through this objective, we can go through all the key aspects of the game, thus making it an intrinsic part of the assessment.

Clearly, the potential for enabling lecturers to incorporate specific items or objects into a student generated game world is an exciting prospect. Not only would this enhance the overall interaction by allowing lecturers to make the game more context specific, but would encourage students to develop new themes of content beyond common industry stereotypes. Linking these together would be the requirement of an overall game objective, which would enable lecturers to navigate through all the essential parts of game world during the assessment.

5. Discussion and recommendations

The research conducted in this paper draws attention to a number of key recommendations and areas for future development for assessments in computer games courses. Despite being derived from a small sample, the findings provide a detailed insight into the intricacies of developing an assessment structure suitable for the unique challenges presented by the emerging field of games design.

Overall, feedback on the assessment structure was positive. Students were generally satisfied with the methods of assessment: an encouraging finding given this was the first set of graduates among a small cohort. The critical requirement was the need to assure students that they will be given appropriate credit during the assessments for specialising in a particular aspect of game development, as is necessary due to the nature and demands of group projects in this field. This requirement was effectively met by the weekly formative assessment/supervisory sessions, which enabled lecturers to gain a much greater understanding of the projects throughout the entire duration of the academic year, and the participation of the key members of the

lecturing team in each viva. Hence, lecturers were able to appreciate the background and development work students had been engaged in, and give this due consideration during the summative assessment stages. Practically, the most significant problem resides with the foreseeable impact on resources should there be substantial increases in student numbers. We predict that the current structure may be efficiently maintained for cohorts of up to around 50 students. In institutions where numbers are greater, changes to viva lengths and/or structure are likely to be required. Based on feedback, there is a strong recommendation that the tiered viva structure should be retained to give students a fair opportunity to explain and demonstrate their work throughout the year, yet this will be (inevitably) challenged by excessive time and resource requirements for large cohorts. As in other disciplines and assessments requiring high levels of lecturer input, the difficulty lies in the provision of a comprehensive assessment structure which reflects the needs of learning outcomes through which students are supported by the necessary supervision, while simultaneously avoiding excessive workload on staff and/or resources. Due to the specific demands of games development and the clear need for viva type assessments, this trade off is likely to become especially prevalent in the future.

In other fields where the resourcing of oral vivas is problematic, one obvious remedy is replacing vivas with written examinations in order to achieve time economies. However, while these may relieve resource issues to an extent, they often present a new set of difficulties which cannot be addressed in the absence of vivas. In this particular case, where students are engaged in year-long projects to produce rich interactive products (beyond that of testing knowledge and the attainment of facts), the provision of the appropriate examination where students are given the opportunity to explain and demonstrate their work is made all the more critical. Indeed, the beneficial nature of vivas can be observed in courses (mostly those in the medical field) where the demonstration of understanding and application of practical skills represents a major component in gauging students' achievement of learning outcomes (see, for example, Shallaly & Ali, 2004; Winning Lim & Townsend, 2005). Thus, while it is impossible to provide a definitive answer for resolving resourcing issues, two likely alternatives include independent lecturer assessments and parallel assessments. The former, as in traditional marking procedures, removes the need for fixed vivas, allowing lecturers to assess the work in their own time checking only for specific learning outcomes. While time economies may be obtained in this approach, the benefits derived from oral presentations will be mostly lost. The latter method retains the viva structure and its benefits, but deploys staff across parallel assessments to reduce time requirements for presentations, and is thus a more conceivable and (more importantly) sustainable solution to significant rises in student numbers.

In contrast, the most exciting opportunity presented in this case study lies in the embodiment of staff and students as virtual characters in the game on which students are being assessed. The feedback from lecturers and students draw attention to a number of contextual and implementation issues which require serious consideration for future assessments of this type. The first critical aspect is that of interactivity, and its essential role in assessments for this discipline. In addition to allowing staff and students to fully engage in the work produced, the use of virtual characters helped to strengthen the contextualisation of game content. There are, however, caveats to their use. One of the main areas of concern is the potential loss of critical awareness on the part of lecturers during the assessment. In the case of *Ertha*, the richness of the interaction and true to life representations of the lecturers led to a highly immersive

and captivating experience. Potentially, this opens the possibility where lecturers' objectivity is reduced by their sheer enjoyment of the game. It is vital to ensure that assessors' perception of the game is not unduly influenced by its visual attractiveness or the novelty of the experience (as these are generally two of the greatest strengths of commercial, interactive games), but to remain solely focused on the evaluation of students' performance in relation to learning outcomes. The other salient recommendation for future projects is that lecturers should have some degree of input into the game and the type of interaction being portrayed. Not only would this offer a level of control over potentially sensitive and/or inappropriate game content, but also a more active opportunity for lecturers to encourage students to develop content across broader themes beyond those routinely developed by the industry. Overall, the likely benefits offered by the use of virtual characters enable much richer assessments to be made, enhancing the experiences of both assessors and students not only on games degrees, but also in related courses where interactivity and the demonstration of visual, digital artefacts are of primary importance.

Although the popularity of computer and video games degrees is steadily rising, many aspects of these new courses are still in formative stages of development. Central to this is the establishment of an appropriate assessment structure through which student achievement and performance may be accurately measured. The task of ensuring that academic standards and procedures meet these requirements is made all the more difficult by the lightning pace at which the games industry develops, necessitating constant changes in technology and working practices which, in turn, alter the type and complexity of content that students are able to produce. It is hoped that the results and recommendations presented in this paper offers practitioners and course developers some insight into how assessments may be conducted for highly interactive, typically bespoke game projects, found on rapidly evolving games courses.

Acknowledgements

The Project Team wishes to thank all the students of the 2006 to 2007 cohort for their enthusiasm and help throughout this research.

References

- Amory, A., Naicker, K., Vincent, J. & Adams, C. (1999) The use of computer games as an educational tool: Identification of appropriate game types and game elements. *British Journal of Educational Technology*, 30(4), 311-321.
- Cagiltay, N. (2007). Teaching software engineering by means of computer-game development: Challenges and opportunities. *British Journal of Educational Technology*, 38(3), 405-415.
- Carless, D. (2002). The 'mini-viva' as a tool to enhance assessment for learning. *Assessment & Evaluation in Higher Education*, 27(4), 353-363.
- Cheung, D. & Siu, B. (2002). Some management issues on computer game development for primary education. *The Electronic Library*, 20(2), 119-124.
- Clark, R. & Choi, S. (2005). Five design principles for experiments on the effects of animated pedagogical agents. *Journal of Educational Computing Research*, 32(3), 209-225.
- Consalvo, M. & Dutton, N. (2006). Game analysis: Developing a methodological toolkit for the qualitative study of games, *Game Studies*, 6(1). [viewed 4 Dec 2007]. http://gamestudies.org/0601/articles/consalvo_dutton

- Davis, M. & Karunathilake, I. (2005). The place of the oral examination in today's assessment systems. *Medical Teacher*, 27(4), 294-297.
- De Aguilera, M. & Mendiz, A. (2003). Video games and education: Education in the face of a "parallel school". *ACM Computers in Entertainment*, 1(1), 1-14.
- Develop (2007). Isn't it all academic? *Develop*, Issue 69, 6-7. Hertfordshire: Intent Media.
- EDGE (2006). *Get into games: Your guide to working in the videogame industry. Supplement*. Bath: Future Publishing.
- EDGE (2007). *Get into games: Your guide to working in the videogame industry. Supplement*, Bath: Future Publishing.
- ELSPA (2007). Computer and video games industry age ratings and codes of practice. <http://www.elspa.com/rating/> [viewed 4 Dec 2007, verified 1 Feb 2009]
- Elwood, J. & Klenowski, V. (2002). Creating communities of shared practice: The challenges of assessment use in learning and teaching, *Assessment & Evaluation in Higher Education*, 27(3), 243-256.
- Feinstein, A., Mann, S. & Corsun, D. (2002). Charting the experiential territory: Clarifying definitions and uses of computer simulation, games, and role play. *Journal of Management Development*, 21(10), 732-744.
- Fry, H., Ketteridge, S. & Marshall, S. (1999). *A handbook for teaching and learning in higher education*. London: Kogan Page.
- Guest, K. E. & Murphy, D. S. (2000). In support of memory retention: A cooperative oral final exam. *Education*, 121(2), 350-354.
- Gulz, A. & Haake, M. (2006). Design of animated pedagogical agents – a look at their look. *International Journal of Human-Computer Studies*, 64, 322-339.
- Hansen, S. (2004). A constructivist approach to project assessment. *European Journal of Engineering Education*, 29(2), 211-220.
- Hasson, F. (2007). Quoted in: Isn't it all academic? *Develop* (Issue 69). Hertfordshire: Intent Media, 7.
- Hong, J-C. & Liu, M-C. (2003). A study on thinking strategy between experts and novices of computer games. *Computers in Human Behavior*, 19, 245-258.
- IGDA (2003). *IGDA curriculum framework. The study of games and game development*. http://www.igda.org/academia/curriculum_framework.php [viewed 4 Dec 2007].
- Inglis, A., Ling, P. & Joosten, V. (2002). *Delivering digitally: Managing the transition to the knowledge media* (2nd Ed). London: Kogan Page.
- Ip, B. & Jacobs, G. (2004). Quantifying games design. *Design Studies*, 25(6), 607-624.
- Jeffery, M. (2007). The war for talent. *Develop*, Issue 69, 49-50. Hertfordshire: Intent Media.
- Kim, Y., Baylor, A. & Shen, E. (2007). Pedagogical agents as learning companions: The impact of agent emotion and gender. *Journal of Computer Assisted Learning*, 23, 220-234.
- Langan, A., Wheeler, C., Shaw, E., Haines, B., Cullen, W., Boyle, J., Penney, D., Oldekop, J., Ashcroft, C., Lockey, L. & Preziosi, R. (2005). Peer assessment of oral presentations: Effects of student gender, university affiliation and participation in the development of assessment criteria. *Assessment & Evaluation in Higher Education*, 30(1), 21-34.

- Laurillard, D. (2002). *Rethinking university teaching: A conversational framework for the effective use of learning technologies*. London; New York: RoutledgeFalmer.
- Maier, P. & Warren, A. (2000). *Integrating technology in learning and teaching: A practical guide for educators*. London: Kogan Page.
- Malliet, S. (2006). Adapting the principles of ludology to the method of videogame content analysis. *Game Studies*, 7(1). [viewed 4 Dec 2007, verified 1 Feb 2009]
<http://gamestudies.org/0701/articles/malliet>
- Malouf, D. (1987). The effect of instructional computer games on continuing student motivation. *The Journal of Special Education*, 21(4), 27-38.
- McFarlane, A., Sparrowhawk, A. & Heald, Y. (2002). *Report on the educational use of games*. TEEM. [viewed 4 Dec 2007, verified 1 Feb 2009].
http://www.teem.org.uk/publications/teem_gamesined_full.pdf
- Moreno, R. (2005). Multimedia learning with animated pedagogical agents. In R. Mayer (Ed.), *The Cambridge handbook of multimedia learning*, Cambridge: Cambridge University Press.
- Moundridou, M. & Virvou, M. (2002). Evaluating the persona effect of an interface agent in a tutoring system. *Journal of Computer Assisted Learning*, 18, 253-261.
- Oyen, A-S. & Bebkko, J. (1996). The effects of computer games and lesson contexts on children's mnemonic strategies. *Journal of Experimental Child Psychology*, 62, 173-189.
- Phillips, R. (1997). *The developer's handbook to interactive multimedia: A practical guide for educational applications*. London: Kogan Page.
- Pillay, H., Brownlee, J. & Wilss, L. (1999). Cognition and recreational computer games: Implications for educational technology. *Journal of Research on Computing in Education*, 32(1), 203-216.
- Prensky, M. (2001). *Digital game-based learning*, New York: McGraw-Hill.
- Rai, S., Wong, K. & Cole, P. (2006). Game construction as a learning tool. *ACM International Conference Proceeding Series: Archive Proceedings of the 2006 International Conference on Game Research and Development*, 223, 231-236.
- Rieber, L. (2005). Multimedia learning in games, simulations, and microworlds In R. Mayer, (Ed.), *The Cambridge handbook of multimedia learning*. Cambridge: Cambridge University Press.
- Rosas, R., Nussbaum, M., Cumsille, P., Marianov, V., Correa, M., Flores, P., Grau, V., Lagos, F., Lopez, X., Lopez, V., Rodriguez, P. & Salinas, M. (2003). Beyond Nintendo: Design and assessment of educational video games for first and second grade students. *Computers & Education*, 40, 71-94.
- Sanches, J. D. (2007). Job's a good 'un. *Develop*, Issue 69, 37-46. Hertfordshire: Intent Media.
- Shallaly, G. & Ali, E. (2004). Use of video-projected structured clinical examination (ViPSCE) instead of the traditional oral (viva) examination in the assessment of final year medical students. *Education for Health*, 17(1), 17-26.
- Sperry, F. (2006). Quoted in: Studio profile – EA UK Studios. *Develop*, Issue 65, 24-25, Hertfordshire: Intent Media.
- Sweetser, P. & Wyeth, P. (2005). GameFlow: A model for evaluating player enjoyment in games. *ACM Computers in Entertainment*, 3(3), 1-24.
- Winning, T., Lim, E. & Townsend, G. (2005). Student experiences of assessment in two problem-based dental curricula: Adelaide and Dublin. *Assessment & Evaluation in Higher Education*, 30(5), 489-505.

Appendix

Interview questions for students

Questions about the general structure of the assessment

Responses for questions 1 and 2 were obtained from all 10 students of the cohort.

1. What are your views regarding the overall assessment procedure? For example, was there enough time allocated to you at each viva stage, or do you feel some stages need more or less time than others? Do you feel that the assessments, including aspects such as assessment format, criteria, and project outputs, were appropriate for the course?
2. Any other views you have regarding the course or assessment stages (such as assessment procedure, criteria, supervision) for the improvement of the course are most welcome.

On interactivity and the use of in-game characters

This question was directed only at the three students involved in the development of *Ertha*.

3. What is your opinion on the use of in-game characters modelled as students and lecturers during the Stage 4 assessment?
4. What are your views in relation to the structure in which the games were played during the Stage 4 viva – i.e. in a linear, student-directed fashion?

Interview questions for lecturers

Questions about the general structure of the assessment

1. What are your general views regarding the entire structure of the assessment (e.g. assessment criteria, formative and summative assessments, issues of formality during each viva, viva duration)?
2. Do you think anything requires changing in view of this being such a new course? How do you feel about the duration of 45 minutes per student (which includes 15 minutes of questions and answers) and the impact on this if student numbers were to increase?

Questions about interactivity and the use of in-game characters

3. What are your views regarding the use of in-game characters that were modeled as the students and lecturers?
4. What are your views in relation to the structure in which the games were played during the Stage 4 viva – i.e. in a linear, student-directed fashion?

Author for correspondence: Dr Barry Ip School of Digital Media, Swansea Metropolitan University Mount Pleasant, Swansea SA1 6ED, UK Email: barry.ip@sihe.ac.uk
