Evaluation of an interview simulation for job seekers - applying a holistic evaluation approach for serious games

Citation for published version (APA):

Document status and date:
In preparation: 01/01/2018

Document license:
CC BY-NC-SA

Please check the document version of this publication:
• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher’s website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.
• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license above, please follow the link for the End User Agreement:
https://www.ou.nl/taverne-agreement

Take down policy
If you believe that this document breaches copyright please contact us at:
pure-support@ou.nl
providing details and we will investigate your claim.

Downloaded from https://research.ou.nl/ on date: 24 Aug. 2021
EVALUATION OF AN INTERVIEW SIMULATION FOR JOB SEEKERS – APPLYING A HOLISTIC EVALUATION APPROACH FOR SERIOUS GAMES

C.M. Steiner¹, A. Nussbaumer¹, K. Gaisbachgrabner¹, T. Platon², O. Lepoivre³, D. Albert¹

¹Graz University of Technology (AUSTRIA)
²BIP MEDIA (FRANCE)
³RANDSTAD (FRANCE)

Abstract

Digital games constitute a major emerging technology that is increasingly entering mainstream educational use. To take full advantage of the great potential of serious games, sound evaluation methods and real-world testing through deployment in authentic educational environments are needed. We present empirical evaluation work conducted in the context of an interview simulation for job seekers. Evaluations have been carried out based on an underlying evaluation framework incorporating the perspectives of development and application of serious games, aiming at investigating the quality and significance of game components used to support game development, the relevance and pedagogical value of a serious game for training, and learning effectiveness as well as learners’ enjoyment of a game. Game developers, training providers and end-users have been involved in the evaluation studies. Results argue for the usefulness of the components applied for game development, the relevance and pedagogical effectiveness of the game application for training, and positive end-user experience while playing the game.

Keywords: Serious games, applied games, evaluation.

1 INTRODUCTION

Digital games constitute a hot topic in educational research for quite some years now and the use of games, gamification and gaming strategies to improve learning is still among the 20 most frequent current topics in e-learning research [1]. Game-based learning has been considered as a major emerging technology in teaching and learning in the NMC Horizon Reports from 2011 to 2015 and meanwhile can be considered a technology that is increasingly entering mainstream educational use.

Serious games (also: applied games, learning games) have a highly engaging and motivating character and enable immersive, meaningful and situated learning experiences [2]. Games can be very effective for learning especially because they may induce a ‘flow experience’ – a positively perceived experience and state of full immersion in an activity that typically goes along with a loss of sense of time [3]. To take full advantage of the great potential of serious games a careful assessment of their quality and impact is necessary. Consequently, there is a need for sound evaluation methods of serious games and for real-world testing of games via deployment in authentic educational environments.

The market potential of serious games is not yet fully exploited. Reasons for that are, among others, the challenge of finding and providing an appropriate balance between gaming and learning and, especially, the high effort and extensive costs required for the creation of successful serious games.

The European project RAGE (http://rageproject.eu/) aims at enhancing the development and take-up of digital game-based learning for game industry and education by providing advanced game components – in order to enable easier, faster and more cost-effective development of serious games [4]. RAGE has developed a collection of more than 40 reusable, and interoperable components for game development. These are made available via an online Ecosystem as a central access point and social space that provides also access to a broad range of literature and training material on serious games. The RAGE game components provide functionalities to undertake various kinds of data analysis, like competence modelling and assessment, emotion detection, or comprehension measurement. Another group of components supports game intelligence and adaptation, e.g. in terms of competence-based personalisation, natural language processing, or social gamification.
In the context of RAGE, a holistic and multi-perspective framework for serious game evaluation has been elaborated that enables the systematic investigation of the RAGE components and games and their effectiveness [5]. In this paper we present how the evaluation framework has been translated into evaluation practice in an educational scenario with real-world users, by the example of a use case and game application for job seekers. The remainder of the paper is structured as follows: section 2 summarises the RAGE evaluation framework and section 3 gives an overview of the application scenario and serious game developed for this purpose. Then, section 4 outlines the evaluation methodology translating the evaluation framework into a research design for comprehensively investigating the development and application of the respective serious game. Section 5 presents the empirical results obtained for the different evaluation dimensions. Section 6 provides an overall summary and conclusion.

2 THE RAGE EVALUATION FRAMEWORK

The overarching goal in the formulation of the RAGE evaluation framework was to gather evidence for the effectiveness of serious game technologies in a systematic and methodologically sound way [5]. The evaluation framework incorporates the perspectives of the different stakeholder groups of serious games in order to ensure (a) the added value and significance of game components for game developers, (b) the relevance and pedagogical value of a serious game for training, and (c) the enjoyment of the game as well as learning effectiveness for learners. (Note: in a more comprehensive version of the evaluation framework also the evaluation of the Ecosystem, which provides access to the game components and other serious games material, is incorporated – this is skipped for the purpose of this paper, as it is not relevant for the empirical work presented here.)

Fig. 1 presents a graphical overview of the evaluation framework. The evaluation framework covers (1) game development and (2) game application for training as the two main dimensions of evaluation. The dimension of game application includes the perspective of training providers as well as end-users (learners). On both evaluation dimensions a levelled approach is taken to represent the relevant evaluation questions and variables.

![Figure 1. RAGE evaluation framework.](image)

Evaluation of game development refers to the effectiveness of component-based game development (i.e. game development making use of RAGE game components). The first level of evaluation relates to the reaction and perception of game developers in terms of usability and usefulness of game components.
components. Level two addresses the relevance of game components (and the functionalities provided) for current or future game projects, thus relating to reasons for using and the intention to use game components. The third level focuses on the analysis of the actual impact of component usage on game development, i.e. the experiences made in technical integration from a game engineering perspective. Level four targets the added value and cost-effectiveness for the game development process.

Evaluation of the game application dimension addresses the educational effectiveness of serious games. To this end, Kirkpatrick’s [6] idea of an integral evaluation of training programmes in terms of a four level process was taken up and adapted. On the first level, learners’ subjective reactions to a serious game in terms of usability and game experience (integrating aspects like enjoyment, usefulness, or flow) are targeted. Level two relates to effects of a game on learning. On this level, it shall be investigated whether and to what extent learning gains in the targeted knowledge/skill domain can be identified after interacting with a game. Level three addresses the question whether learners are able to apply (i.e. transfer) the knowledge and skills acquired during gaming to other contexts, in particular to real-world settings. Level four is dedicated to a more global and organisational perspective on the pedagogical value of serious games for training providers and/or educational institutions, in particular the costs and benefits for introducing and using this type of learning technology for training.

The evaluation framework served as a common ground for all evaluation work within the RAGE project, with shared methodologies across different application scenarios and games, while nevertheless providing flexibility to accommodate to the specific conditions of the individual application scenarios. An iterative evaluation approach was pursued implementing formative and summative evaluation phases aligned with the technology development phases. With regards to evaluation instruments, the framework calls for a mixed-method approach in evaluations, enabling the integration and triangulation of qualitative and quantitative data from multiple sources and perspectives.

3 THE JOB SEARCH USE CASE AND GAME

The RAGE technologies and methodologies were applied and tested in the context of serious games for six selected application scenarios representing different target groups and markets. For each use case a game was created using and integrating a subset of RAGE components in game development.

One application scenario was in the context of recruitment training and services. Concretely, the use case of a major international employment and recruitment agency that provides education and training for recruiters and temporary workers was used. To address and reach a wide variety of target groups, new methods for teaching and training, particularly in the area of employability skills are being sought by the recruitment agency. One key issue identified in terms of training needs is the lack of appropriate job search competences that many job seekers have. The goal was to develop a serious game for job search skills training, to be offered to candidates as support tool for preparing their job application and recruitment interview.

A game called 'Job Quest' was developed for the use case. It is a single player game simulation of a real job search experience and aims at supporting end-users in handling their job searches, in particular job interviews (see Fig. 2 for a screenshot). The targeted learning objective is to convey job search skills. In the game, different job offers are presented to the player from which to choose from to undergo a job interview. Afterwards, the player gets feedback and advice on interview performance. A range of interviews can be done, providing the experience that different kinds of recruiters will use slightly different interview approaches with different objectives. In addition, the game identifies a players’ interest profile (for adapting the game to the player) and provides an optional CV analysis with subsequent advice on CV appearance and content. For Job Quest several RAGE components have been used – some to accelerate and facilitate game development (e.g. authentication, interaction tracking and analytics) and some others to add unique functionalities to the game (e.g. real-time emotion detection, text analysis). The game was deployed and evaluated at several branches of the employment and recruitment agency.
4 EVALUATION METHODOLOGY

Using the evaluation framework presented in section 2 as a reference point, evaluations of the RAGE game components and of the application of the Job Quest game for training were conducted. Here we present the results from the final evaluation round on both evaluation dimensions.

4.1 Game Development

Evaluation on the game development dimension aimed at investigating the quality and added of the game components for developing serious games. Feedback on the evaluation levels reaction (in terms of usability of components and their perceived usefulness), relevance, game engineering, and costs and benefits for game development, as presented in Fig. 1, was collected from developers who have applied RAGE components in their game development (including, among others, the Job Quest game). Each participant gave feedback exclusively on those components that he/she had made use of.

An online survey was used as evaluation instrument. The survey gathered feedback on the evaluation topics through closed-ended Likert-type items and open-ended questions. A combination of standard questionnaires and survey items specifically defined for the purpose of this evaluation was used.

- **Reaction:** For usability assessment the UMUX questionnaire (usability metric for user experience [7]) covering 4 items was used. Usefulness was assessed by adopting 4 items from the ‘perceived usefulness’ scale of Davis’ instrument for user acceptance ([8], [9]).
- **Relevance:** The relevance of game components for game projects was assessed with one Likert-item and two open questions eliciting the further detail on the reasons for choosing and using a component and intention for future use.
- **Game engineering:** This more technical point of view addressing component integration in the usual/traditional game engineering process was captured by an open question.
- **Game development costs and benefits:** Experienced or expected benefits were captured by two Likert-items and an open question. Cost effectiveness of using and integrating a component in game development was assessed via four Likert-items.

4.2 Game Application

The evaluation of the application of Job Quest for training aimed at investigating the educational effectiveness of the game in line with the evaluation levels presented in Fig. 1. Evaluation data and feedback was gathered from end-users, i.e. job seeking candidates involved at the different branches of the recruitment agency, on reaction (usability and game experience, with game experience represented by the aspects enjoyment and usefulness) and learning. On the learning level the learning effectiveness of the game to support job search skill development was examined. Job search efficacy was operationalised in terms of job-seeking self-efficacy, which can be understood as the perceived
ability to perform the skills involved in seeking employment. (Note: An explicit evaluation of transfer, i.e. evaluation level 3, was left out for the purpose of this evaluation study.) Pedagogical costs and benefits of using the serious game were evaluated with training providers represented by consultants from the recruitment agency.

4.2.1 End-users

For the evaluation of reaction and learning, a pre-post-test design was used (see Fig. 3); evaluation data was gathered via questionnaires before and after playing the game, to analyse the effect of playing the game on job search efficacy and to analyse users’ game experience. The pre-test collected demographic information and prior job search experience and included an assessment of job-seeking self-efficacy as baseline measure. Then, a game session followed; in the game the player met a virtual personal recruitment advisor and could do up to six job interviews, as well as an optional CV analysis. The post-test phase included usability and game experience measurements and again a self-assessment of job search efficacy to examine learning effects due to playing the game.

In terms of evaluation instruments, a mixed-method approach combining questionnaires (pre- and post-test), in-game data and verbal qualitative feedback was used.

**Questionnaires:**
- **Demographic information and job search experience** were collected in the pre-test via 6 items.
- **Reaction:** For a general usability assessment in the post-test an item targeting the ease of use of the game was adopted from the usability metric for user experience (UMUX [7]). Game experience was evaluated in terms of enjoyment and usefulness. As a measure of enjoyment of the game 2 items addressing the perception of entertainment vs. boredom and the recommendation of the game to others were used, adapted from the Game User Experience Satisfaction Scale (GUESS [10]; subscale ‘Enjoyment’). Usefulness of the game for job search skills training was assessed via 3 items defined for the purpose of this study and in line with the contents of the game.
- **Learning:** To determine a baseline measure and learning effects with respect to job search efficacy in the pre-test 7 Likert-type items were used, adapting items from the Career Self Efficacy Scale (CSES [11]; subscales ‘job search efficacy’ and ‘interviewing skills’) and the Job-Seeking Self-Efficacy Scale (JSS, [12]). For a follow-up assessment of job-seeking self-efficacy in the post-test a subset of 4 items was presented again as a short version of the self-assessment instrument.

In-game data: For a closer look on game experience, in-game activities (i.e. players’ interactions with the game) were tracked during the game session. To this end, an evaluation component [13] had been integrated in the Job Quest game, which made it possible to pool, store, and analyse game event data for the purpose of evaluation.

**Qualitative feedback:** Comments and question from participants during the introduction of the game and while playing, as well as observations of critical events occurring during the game session were recorded to serve as additional information for the evaluation of usability, game experience, and perceived effectiveness for learning.

---

**Figure 3. Research design for end-users.**
4.2.2 Training Providers

Feedback from training providers on the pedagogical costs and benefits of introducing and using the Job Quest game for training and, respectively, serious games, in general, as a learning technology was gathered via a semi-structured interview questionnaire with 9 open questions.

5 RESULTS

5.1 Game Development Results

In total 14 game developers from game industry and academia took part in the evaluation. To derive in-depth feedback and possible action-points for component developers, survey responses were analysed in detail for each individual game component. For the purpose of this paper overall scores and conclusions obtained and summarized for the entirety of all game components are presented, to give evidence on the quality and effectiveness of RAGE component usage, in general.

From the ratings on the closed-ended survey items scores for each evaluation level were calculated. For usability (in line with the scoring scheme of the UMUX scale) the possible score range was 0-100, with higher values indicating better usability. All other scores had a possible range from 1 to 7, with higher values indicating better results.

As can be seen in Fig. 4, a usability score of 70.36 (SD=19.35) was measured, indicating good overall usability for RAGE components. It has to be noted, though, that there was a quite high variation of assessments for individual components, which is reflected by relatively high standard deviations. Concerning usefulness, an overall score of 5.16 (SD=1.33) was found, which suggests that the components are perceived as useful in game development. The components were assessed with good overall relevance (M=5.11, SD=1.51). Also with respect to perceived benefits for game development and cost effectiveness of component usage positive overall results could be obtained, with an average benefit score of 5 (SD=1.39) and a mean score for cost effectiveness of 4.8 (SD=1.34).

From the qualitative feedback obtained through open questions it became clear that the relevance of using game components was primarily due to the specific functionality provided by the respective software component. Clear intention to use components, also in future game applications, was signalled, provided the applicability and need of the respective functionality for a game project. Concerning game engineering, respondents especially pointed to the possibility of reducing costs, i.e. resources, effort and/or time required for game development. They also pointed to the advantage of using pre-tested code and the ease of integration and modification of components. Feedback on the benefits for game development and serious games resembled the usefulness and significance of the functionalities provided by the game components, the possibility to add unique functionality, but also the benefits in terms of making the game production process faster and easier.
5.2 Game Application Results

5.2.1 End-user Experience

In total 369 candidates (51% female, 35% male, 14% did not specify) tested and played the Job Quest game. A subset of this overall sample also completed the evaluation questionnaires. After data cleaning, responses of 291 participants for the pre-test and 78 participants for the post-test were analysed in more detail. Participants were on average 33 years (SD=11). Slightly more than half of the participants (51%) had previous experience in recruitment interviews.

Reaction

Analysis of in-game data showed that a mean number of 3 interviews (SD=0.58) was done by players during gaming, out of 6 on principle possible interviews. The extent to which players were active in the game and explored the possibilities of the gameplay gives some indication of users’ game experience. An intensive interaction with the game and high number of accomplished game missions (i.e. interviews) may be considered as indicating a positive engagement with and experience of the game. The tracked player interactions with the game may be considered as reflecting satisfying game experience.

For the responses on usability (M=1.4), enjoyment (M=1.5) and usefulness (M=1.5) in the post-test, a strong ground effect was identified, which at first sight seems to reflect quite critical player opinions. When contrasting these results with the verbal qualitative feedback during game sessions, though, a totally different image emerges, such that the questionnaire scores need to be attributed to inaccurate and careless answer behaviour rather than to poor user experience. Indeed, qualitative feedback clearly showed that participants were globally very interested in the game and that they found the game provided nice user experience. The advice that the game provides for supporting and improving job seeking was very positively perceived. The qualitative feedback therefore gives indication of a satisfactory usability of the game and of good user experience.

Learning

For job search efficacy an overall score was calculated, with a possible score range from 1 to 5. The mean score obtained from the pre-test was 3.0, which indicates moderate self-assessed job search knowledge and skills. Variation between individual participants was quite high (SD=1.54). When analysing the relationship between job search efficacy and experience in recruitment interviews, a significant correlation (r(291)=.75, p≤.001) could be found, indicating that participants with prior job interview experience assessed their job search skills better than participants who have never had a recruitment interview.

Although no statistically significant learning effects could be identified when comparing the self-assessment of job search efficacy before and after playing the Job Quest game, the qualitative feedback received from participants underlines that they appreciated having the opportunity of this kind of service/training. This is further encouraged by feedback from training providers, highlighting that the game was helpful for job interview preparation. In particular, one recruitment consultant even highlighted an explicit learning effect, i.e. that an improved job search self-efficacy could be noticed in consulting with candidates after they had played Job Quest, compared to candidates who did not use the game.

5.2.2 Training Provider Feedback

Pedagogical Cost and Benefits

Feedback on pedagogical costs and benefits was elicited from five recruitment consultants. The main benefits of using the game for training were seen in the innovative and progressive approach to training, the entertaining character of the game, and the easy and quick applicability of the game for training. The game was perceived as reducing the extent of face-to-face training and consultancy needed. Besides, it was considered as a new opportunity of establishing contact with candidates or as a suitable tool for follow-up support. Additional costs mentioned related to additional effort and time required for introducing the game, the need of recruiters as well as candidates to get used to the game application, and the time required to assist candidates, to provide explanation and support. These were, however, assumed to be balanced by the fact that face-to-face consultancy was more effective and required less time after candidates had used the game. The deployment of a web-based version of the game was considered an aspect for further improvement that would further enhance cost-benefit balance and broaden the game’s applicability. Overall, respondents agreed that the use of
serious games is a worthwhile investment for the organisation, and that the innovative image that can be conveyed by the use of this type of learning technology was highlighted as a main competitive gain.

6 OVERALL SUMMARY AND CONCLUSION

The RAGE project has developed and made available a whole range of game technology components and knowledge resources aimed at supporting game developers and game studios in developing serious games more efficiently and making them better suited for their purpose. The game components have been integrated and tested in game projects for a set of selected use cases, the resulting games were deployed in real-world training settings. As a basis for comprehensively evaluating the different project and technology outputs, an evaluation framework has been elaborated incorporating the dimensions and stakeholders of game development and game application in training. This multi-perspective approach provides a basis for establishing a holistic understanding of the quality and impact of the serious game technologies, from the component-based game development process to the actual interactions with and impact of the resulting serious games.

We presented the evaluation framework and how it was translated into and used in evaluation practice. The ‘Job Quest’ game was developed using a subset of RAGE game components and is a game simulation of a real job search experience, to serve as a training approach supporting face-to-face recruitment training in the human resource services industry. Empirical evaluations were carried out incorporating the evaluation questions from the different stakeholders’ perspectives. The evaluation results give evidence of the relevance and significance of the RAGE components for serious game development, thus arguing for the continuity and sustainability of these technologies. The evaluation of the application of the Job Quest game in training showed that game experiences obtained positive feedback, and that the games’ relevance in the context of recruitment training and its potential to support learning was well recognised.

Overall, the applied evaluation methodology proved as a suitable and holistic approach incorporating the evaluation interests and needs from different perspectives. The benefits of the RAGE approach and technologies for game development and educational application could be demonstrated and useful information and inspiration for future work could be obtained.

ACKNOWLEDGEMENTS

This work has been partially funded by the EC H2020 project RAGE (Realising and Applied Gaming Eco-System); http://www.rageproject.eu/; Grant agreement No 644187. This document reflects only the views of the authors and the European Commission is not responsible for any use that may be made of the information it contains.

REFERENCES


