Learning analytics in serious gaming: uncovering the hidden treasury of game log files

Open Universiteit
www.ou.nl

Learning analytics in serious gaming: uncovering the hidden treasury of game log files

Wim Westera[1], Rob Nadolski[1] & Hans Hummel[1]

[1]Open University of the Netherlands, Valkenburgerweg 177, 6419 AT Heerlen,
The Netherlands
{wim.westera, rob.nadolski, hans.hummel}@ou.nl

Abstract. This paper presents an exploratory analysis of existing log files of the VIBOA environmental policy games at Utrecht University. For reasons of statistical power we've combined student cohorts 2008, 2009, 2010, and 2011, which led to a sample size of 118 students. The VIBOA games are inquiry-based games, which offer a lot of freedom of movement. Our premise is that this freedom of movement is accompanied by behavioural variability across individuals, which may influence the efficiency of learning. Descriptive statistics of our sample revealed such variability of diverse game parameters. We have identified "switching behaviour", defined as the number of game objects (videos, resources, locations) accessed per unit time, as a relevant behavioural pattern. Multiple regression analysis showed that switching rates of videos and locations explain 54 % of the variance of learning efficiency (defined as final score per unit time). Both the model and the model coefficients were significant beyond the 0.001 level. The same switching variables also account for 45% of the variance of total time spent T. Predictive models of final score weren't found. We conclude the paper by critically evaluating our findings, making explicit the limitations of our study and making suggestions for future research that links learning analytics and serious gaming.

1 Introduction

Serious games are outstanding examples of adaptive systems as they continuously adjust their responses to the learners' actions for preserving favourable game play. Although many serious games retain somewhat of the user's history in log files, the actual use of logging data is quite limited. Adaptive game responses seldom build on the user's history of consecutive actions, but instead tend to use a static set of criteria, very similar to a simple checklist. Similarly the assessment of learning progress is based on closures and performance milestones. From a game developer's perspective this situation is understandable, since the only relevant thing would be to check whether the player achieves sufficient performance milestones within the constraints

of the game rules. However, from an educator's perspective, if not a research perspective, the players' logging history could be beneficial for building detailed user models, analysing the process of learning or tracing bottlenecks in game play. For preserving the efficiency of learning it is highly relevant to gain insights in the activities and behaviours needed for the player to reach these performance milestones: e.g. did the learner achieve the milestone in an efficient and well-considered way, or was it a thoughtless trial and error style that took a lot of time without achieving any learning gains? Various authors [1,2] explain the difference between a performance orientation and a learning orientation: while game play tends to focus on performance, which is linked with an attitude of achieving milestones and score (in many cases under time constraints), learning requires opportunities for reflection, repetition, self-evaluation, pauses, and even the preparedness to make mistakes. Hence, the process of gaming may readily counteract the process of learning. Having completed a serious game successfully doesn´t necessarily imply successful learning. This uncertainty about the effectiveness of learning will be larger as games offer more freedom of movement to the learners. In well-structured drill and practice games, e.g. in arithmetic, learning gains are likely to coincide with performance gains. In recent years, however, serious games have been based more and more on open approaches associated with contextualized problem solving, adventure games, inquiry-based learning and competence learning. These approaches all offer large degrees of freedom and heavily rely on self-directed learning, self-regulation and other 21st century skills relevant for today's knowledge workers [3,4]. In such contexts the freedom of movement is likely to display more behavioural variability across different individuals. So far, however, very little is known about the actual in-game behaviours of learners and how these behaviours relate to learning achievements. Game logging data are a treasury of information, which could be exploited for distilling more details about the players' learning achievements based on their wanderings and trajectories through the network of game state nodes.

This paper provides an exploratory study of existing log files of serious games that were used by 118 master students in environmental sciences students at Utrecht University. For practical reasons the study is constrained to a retrospect study: the logging files date back to 2008-2011. Our aim is to explore to what extent the logging data of these particular games would be helpful to reveal meaningful patterns, variables and relationships. We will first briefly explain the emerging research field of learning analytics and connect this with new developments of user tracking in serious games. We will describe our serious games and their context of use. Next we will describe the type of logging data that are available and explain what primary variables will be used for the analysis. Finally the results are presented and critically evaluated.

2   Learning analytics

Analysing learner log data is closely related to the wider framework of learning analytics, which is generally defined as using the ever-growing amounts of data about learners' activities and interests for improving learning outcomes [5]. A related term

is educational data mining [6]. While the focus of educational data mining is on methods for extracting the data, learning analytics concerns the development and application of predictive models in instructional systems [7]. Although student metrics have been used for decades to improve student throughputs and education workflows, in recent years datasets have grown larger and have become more easily accessible because of the use of Learning Management Systems such as Moodle and Blackboard for the online delivery of learning content. Many of these tools include student tracking functions that automatically log and aggregate student activity data as well as user profile data, access statistics and test results. At an institutional level a best practice is provided by the Signals project at Purdue University, which demonstrates how educational data mining and predictive modelling can be used for obtaining higher grades and retention rates than were observed in control groups [8,9]. New developments such as open educational resources and MOOCS [10] will yield big sets of learner data and support their analysis. Although learning analytics is generally qualified as an opportunity for improving the quality and effectiveness of learning, important concerns are raised because analytics could severely disempower and demotivate learners when they are provided with continuous feedback about their knowledge and performance gaps as compared with other students [11]. Also the capturing of unstructured personal traces across different platforms, social networks and contexts goes with some principle barriers linked with privacy protection and other legal issues [5].

3 Serious gaming and logging

As Gee [12] pointed out many video games are based on the growth of the players' mastery of knowledge and skills, and therefore incorporate tracking methods that allows for continuous adaptation of difficulty levels, hints and challenges to their achievements and progress. Some games provide learner support services based on playing behaviour, e.g. scaffolding, hinting, micro-feedback, meta-level feedback [13,14]. Social gaming companies, such as Zynga, use player tracking for predicting what users want and will do next in a game to find out how to make games more fun and get users more engaged [7]. In most serious games, however, the logging is mainly used for triggering events and new episodes in the game flow and game narrative, but is seldom used for the accommodation of user modelling and personalised learning. Although user tracking is a predominant and well-exploited mechanism for adaptive gameplay, the player's full history of states is greatly underused, because most games reflect a discrete time Markov chain, which assigns only a limited role to state history and process memory. Serrano-Laguna et al. [15] notice that logging data in games are highly under-exploited. They suggest that logging of interactions in a game could in principle be used for automated assessment as an alternative for intermediate tests or questionnaires, which are often perceived as unwanted interruptions of game play. A comprehensive approach for in-game assessment was given by Shute et al. [16], which entails the combination of Evidence-Centered Assessment [17] and

Bayesian score models [18]. Its ingredients include a competency model, a learner model, an evidence model (providing clues for evidence) and a task model. Reese and colleagues [19] report about the CyGaMEs project, which quantifies game play activity to track timed progress toward the game's goal and uses this progress as a measure of player learning. Westera et al. [4] point at using game logging data for strategic feedback rather than tactic or micro-level feedback, in order to nourish the learners' self-directedness, self-evaluation and reflection.

4　Serious gaming context: Environmental policy games for a Master of Science degree

Our analysis focuses on a series of 5 online serious games in the domain of environmental sciences that were jointly developed by the Open University of the Netherlands, Utrecht University, and Radboud University Nijmegen. The games are implemented on the EMERGO game engine (www.emergo.cc) and linked together in a single game run that takes about 50 hours of study. Students preparing for their Master of Science degree adopt the role of an environmental consultancy trainee at the (fictitious) VIBOA consultancy agency. In this role they are confronted with authentic environmental policy problem cases (1. Wadden Sea, 2. Wind energy, 3. Lake Naarden, 4. Micro pollution, 5. River management). Guided by inquiry-based gaming scenarios the students individually have to make a thorough analysis of the problems by consulting various stakeholders (video-based interviews with experts) and by collecting and combining relevant information from reports, scientific papers, texts of law, formal documents and other sources. They have to apply scientific methodologies and theories, and finally deliver a report for proposing well-substantiated policy measures. The game narrative and feedback is closely linked with the work conditions at the environmental consultancy office. In figure 1 a screenshot of the online game shows a scene from a (videotaped) meeting at the VIBOA environmental consultancy office.

Fig. 1. Screenshot of the VIBOA environmental consultancy game

The games include 4 pre-tests which are used for self-testing of prior knowledge before entering a game. Each of the games offers large degrees of freedom to the students as to what approach to develop, who to interview, or what sources to use. Occasionally, incoming notifications or (pseudo) email messages provide new information, announce new events, provide hints or prompt for certain actions. The first game is an introductory game. The examination is based on the students' reports about games 2, 3, 4 and 5.

We've retrieved logging data of student cohorts from Utrecht University. For reasons of statistical power we have combined cohorts of 2008, 2009, 2010, and 2011, all playing the games under the same internal and external conditions (e.g. preparation, time frames, intermediate assessments, examinations). After exclusion of students who failed to obtain a final score (e.g. dropouts), we ended up with a sample of 118 students (46% male, 54% female) including 7 people who failed in one year but re-enrolled in the subsequent year.

5   Theoretical framing

The observations that we will be able to make are constrained by the type of logging data that we have available. The EMERGO educational game engine is an open source web-based game authoring and run-time delivery environment composed of diverse functional components. These components are linked together by the game logic. Since each component has its own separate logging system, we have implemented a logging aggregator, which integrates the distributed data into a joint logging file. The logging file records all meaningful student actions such as visiting a location, opening an information resource (document, URL, picture, graph, etc.), accessing a video, asking a video-interview question, accessing a pre-test, answering a pre-test

question and some more. Also system responses are recorded. All actions go with a timestamp. Figure 2 displays a specimen of the aggregated log file.

name	time	component	tag	id	key	value	system/user	userdata	data
Student XXX	21436	script Windenergie	timer	TI_Start Deeltoets 2	finished	true	system	none	none
Student XXX	21439	gesprekken	map	theoretisch kader en onderzoekmethode	outfolded	true	user	none	none
Student XXX	21454	gesprekken	question	Vanuit welk theoretisch perspectief op belei	opened	true	user	none	none
Student XXX	21454	gesprekken	question	Vanuit welk theoretisch perspectief op belei	outfoldable	true	system	none	none
Student XXX	21454	gesprekken	question	Vanuit welk theoretisch perspectief op belei	selected	true	user	none	none
Student XXX	21454	gesprekken	fragment	OGB-WND-SAG-INT-006	opened	true	system	none	OGB-WND-SAG-INT-006.flv
Student XXX	21457	script Windenergie	timer	TI_Start Deeltoets 2	finished	true	system	none	none
Student XXX	21463	gesprekken	question	Welke methoden of modellen heb je in dit o	opened	true	user	none	none
Student XXX	21463	gesprekken	question	Welke methoden of modellen heb je in dit o	outfoldable	true	system	none	none
Student XXX	21463	gesprekken	question	Welke methoden of modellen heb je in dit o	selected	true	user	none	none
Student XXX	21472	gesprekken	question	Was de toepassing van de methoden en mod	opened	true	user	none	none
Student XXX	21472	gesprekken	question	Was de toepassing van de methoden en mod	outfoldable	true	system	none	none
Student XXX	21472	gesprekken	question	Was de toepassing van de methoden en mod	selected	true	user	none	none
Student XXX	21472	gesprekken	fragment	OGB-WND-SAG-INT-009	opened	true	system	none	OGB-WND-SAG-INT-009.flv
Student XXX	21475	gesprekken	question	Heb je zelf een nieuwe methode of methodi	opened	true	user	none	none
Student XXX	21475	gesprekken	question	Heb je zelf een nieuwe methode of methodi	outfoldable	true	system	none	none
Student XXX	21475	gesprekken	question	Heb je zelf een nieuwe methode of methodi	selected	true	user	none	none
Student XXX	21475	gesprekken	fragment	OGB-WND-SAG-INT-010	opened	true	system	none	OGB-WND-SAG-INT-010.flv
Student XXX	21478	gesprekken	component	component	opened	false	system/user	none	none
Student XXX	21478	gesprekken	conversation	WND INTERVIEW SUSANNE AGTERBOSCH	opened	false	system/user	none	none
Student XXX	21478	gesprekken	conversation	WND INTERVIEW SUSANNE AGTERBOSCH	finished	true	user	none	none
Student XXX	21478	locaties NIEUW	location	LOC-Terug naar Hal	accessible	true	system	none	none
Student XXX	21478	script Windenergie	condition	IF_OGB-WND-SAG-interview_AFGEROND	opened	true	system	none	none
Student XXX	21478	script Windenergie	action	THEN_LOC Terug naar Hal_TOEGANKELIJK	opened	true	system	none	none

Fig. 2. Specimen of the aggregated log file

A brief explanation of the log file: the "time" column displays the timestamps expressed in seconds, the "component" column refers to the EMERGO software component involved, the "id" column refers to a specific authored game content object that is addressed, "key" and "value" represent the action exerted, the "system/user" column refers to the actor (either the user or the system, or both). For instance, at time 21454 the user selects a question to ask to an expert, whereupon the system responds by playing the Flash video file OGB-WAD-SAG-INT-006.flv, which contains the recorded answer of the expert.

From the logging data we were able to derive descriptive statistics (e.g. means, standard deviations, variances) of a wide range of variables. For this first explorative study we have confined ourselves to a basic set of variables that are retrievable for each student with simple queries:

- T — Total time spend
- N_L — Number of locations (re-)visited
- N_R — Number of information resources accessed
- N_V — Number of videos accessed
- N_P — Number of pre-test answers given, including improvements
- S_P — Pre-test score (initial answers only)
- S_F — Final assessment score assigned by the examiner on the basis of submitted reports.

Locations are spaces in the game environment, for instance "rooms" where the players may find specific information or where they may talk to an expert (video interviews). Resources include relevant papers, reports, letters or other documents, URLs, graphs etcetera that are made available in the game. Videos can be any record-

ed file including expert interviews, instructional videos, archived TV-programs or documentaries. The pre-test scores S_P are based on the initial answers of the students to the pre-test items (40 items in total). Since students were allowed to change their answers, we used their initial answers to obtain a metric of prior knowledge. The final assessment scores S_F are the only data not derived from the logging, but they are assigned by the examiner on the basis of the reports the final report about the respective games.

In view of the inquiry-based gaming environment and the freedom of moment for students we focused our study on the following research questions:

- To what extent can we identify different gaming behaviours?
- To what extent can behavioural characteristics be predictors of the final assessment scores?
- To what extent can pre-test scores be predictors of final assessment scores?

6 Results

We used MS-Excel for filtering the log files of 118 students (up to 700,000 records) and used SPSS for statistical processing. Few outliers were traced (in all cases at most two outliers with z-score>3), which were kept in the sample.

6.1 To what extent can we identify different gaming behaviours?

In Table 1 some of the key figures of the logging analysis are summarized.

Table 1. Descriptive statistics of the logging data

	Average per student	Standard deviation	Coefficient of variation
Total time T	53.8 h	24.6 h	0.44
Pre-test score S_P	6.6	1.2	0.18
Final score S_F	6.6	1.5	0.24
Resources accessed N_R	180	96	0.54
Videos accessed N_V	120	42	0.35
Locations accessed N_L	156	61	0.39
Pre-test answers N_P	65	42	0.65

Total time required for the 5 games is about 54 hours. The standard deviation of 24 hours indicates considerable spread among students. Similar large spreads are observed in the amounts of accessed resources, videos, locations and pre-test answers, indicating substantial variability of playing behaviours. Note that the games provide access to a limited set of resources (89), videos (100), locations (23) and pre-test questions (40). So, many of these are revisited. Some locations re-occur in all games and have to be re-opened in each game. It should be noted that the pre-tests and the final test are not equivalent: the pre-tests are a simple check of basic knowledge re-

quired for entering the games, while the final test covers all contents covered by the games.

For analysing the coherence of variability across different behavioural indicators we calculated correlations between the access numbers N_R, N_V and N_L. We omitted pre-test questions N_P because the pre-tests are preceding the actual game play. Table 2 shows the results.

Table 2. Correlations of accessed resources, locations and videos

Variables		Correlation r	p-value
Accessed resources N_R	Accessed locations N_L	.489	<.001
Accessed resources N_R	Accessed videos N_V	.447	<.001
Accessed videos N_V	Accessed locations N_L	.351	<.001
Access rates resources N_R/T	Access rates locations N_L/T	.500	<.001
Access rates resources N_R/T	Access rates videos N_V/T	.484	<.001
Access rates videos N_V/T	Access rates locations N_L/T	.576	<.001

We found moderate correlations of $r_{R,L}=.489$, $r_{R,V}=.447$ and $r_{V,L}=.351$ all at a significance level of p<0.001. In view of the variance of total time T we replaced the access numbers N_R, N_V and N_L with the access rates N_R/T, N_V/T and N_L/T, respectively. Here we found slightly higher correlations of $r_{R/T,L/T}=.500$, $r_{R/T,V/T}=.484$ and $r_{V/T,L/T}=.576$, all of which are likewise highly significant (p<0.001). It demonstrates that all variables point at the same direction, be it that the variability of one rate variable explains between about 23% and 33% (r-squared) of the variability of another rate variable. This signals some behavioural consistency between the access rates of resources, videos and locations. Both the correlation and the variance identify the students´ "switching" behaviours as a likely behavioural trait.

6.2 To what extent can behavioural characteristics be predictors of the final assessment scores?

Given the observed variability of "switching" behaviours we have looked into predictive regression models that use the switching rates N_R/T, N_V/T and N_L/T as predictors of learning gains. Regarding these time-based activity rates we expressed the learning gains as scores per unit time S_F/T (this represents the efficiency of learning). Table 3 shows the results of the multiple regression analysis (hierarchical forced entry).

Table 3. Hierarchical regression analysis of learning efficiency and switching behaviours

Model	Video switching rates N_V/T included	Location switching rates N_L/T included	Resources switching rates N_R/T included	R^2	p-value
Learning efficiency S_F/T	yes	no	no	.430	<.001
Learning efficiency S_F/T	yes	yes	no	.540	<.001
Learning efficiency S_F/T	yes	yes	yes	.545	<.001

We found that the overall model has an explanatory power of R^2=.545 ($F(3,114)$= 45.6, p<0.001) which means that more than half of the variability of scores can be explained by the variability of switching behaviours. In the model hierarchy the adjusted R^2 increased from 0.430 (videos only: $F(1,116)$=87.5, p<0.001) to 0.540 (videos and locations: $F(2, 115)$=27.5, p<0.001) and 0.545 (videos, locations and resources: $F(3,114)$=1.3, p<0.249). It follows that only videos and locations contribute significantly to the model. Regression coefficients in the reduced model are b_1=0.037 ($t(117)$=5.5, p<0.001) for the video rate and b_2=0.022 ($t(117)$=5.2, p<0.001) for the location rate. It means that switching behaviour as based on video access rates and location access rates is a predictor (54%) of learning efficiency.

High learning efficiency, however, isn't equivalent with high final score, because the latter may also depend on total time spent. To what extent is switching behaviour a predictor of total time spent? Do fast switchers study faster? Multiple regression analysis with switching behaviours N_V/T and N_L/T as predictors for total time T produce a significant model with R^2=0.445 ($F(2,115)$=46.0, p<0.001) and standardized coefficients b_1=-0.431 ($t(117)$=-5.1, p<0.001) and b_2==-0.318 ($t(117)$=-3.7, p<0.001). This suggests that the variance of switching behaviours can account for 45% of the variance of total time spent. The negative signs of the regression coefficients indicate that high switching rates correspond with low T. Unfortunately, a combined model of using switching behaviours and total time spent T for predicting final scores S_F failed to produce meaningful outcomes. This is consistent with the very weak correlation that we've observed between scores S_F and total time spent T (r=0.182, p=0.049). Apparently other factors are predominant in final scores.

6.3   To what extent is prior knowledge a predictor of final assessment scores?

Having low levels of prior knowledge means that a learner has to make more efforts for achieving the same learning outcome. Alternatively, making the same learning effort as learners with sufficient prior knowledge would procure lower final scores. This line of reasoning suggests a model with pre-test score and total time as predictors of final score. Also the learning would be proportional with learning efficiency. As has been explained above learning efficiency is partly predicted by switching behaviours. Regressions with the current dataset, however failed to confirm a combined model of switching and pre-test score to predict final score. Also, simplified models using fewer predictors failed to produce meaningful outcomes. We've also used the

number of answers N_P given in the pre-tests, which is possibly an indicator of ignorance or trial and error behaviour, as a predictor. But N_P didn't contribute significantly to the regression statistics.

7 Discussion and conclusion

In this study we analysed existing logging files of the VIBOA environmental policy games at Utrecht University. The freedom of movement that these inquiry-based learning games offer, seems to be accompanied by substantial behavioural variability across different individuals. Descriptive statistics of our sample of 118 subjects revealed such variability of the number of accessed locations N_L, the number of accessed videos N_V, the number of accessed resources N_R, the number of pre-test answers N_P given, and the time T spent to the games. On average the game's (written) resources are opened twice, while videos tend to be opened only once. An explanation might be in the very modality of the objects: written resources are randomly accessible, while video has a temporal nature. It takes time to review a video, while a written resource could be re-accessed easily for looking up things. A technical issue may even have concealed the true rate of resource consultation, because in contrast with videos and locations all resources automatically open in a new browser window, which can be re-consulted by the player without making the game engine aware of this (therefore we didn't make extensive use of the resource variable N_R in our analyses). We found moderate but significant cross-correlations of N_R, N_V and N_L and even higher values between the access rates N_R/T, N_V/T and N_L/T, respectively, suggesting the relevance of a behavioural trait that could be identified as "switching behaviour". Multiple regression analysis showed that a model which uses video access rates N_V/T and location access rates N_L/T as predictors explains 54 % of the variance of learning efficiency. Both the model and the model coefficients were significant beyond the 0.001 level. The sample size of 118 subjects is well above the minimum requirements according to Greens [20] rules of thumb (50+8k, and 104+k with k the number of predictors), which strengthens the reliability of the outcome. We also were able to produce a reliable and significant model of switching behaviours $N_{V/T}$ and $N_{L/T}$ predicting total time T. The silent but reasonable assumption here is that the dependent variable T is not necessarily required for determining the switching rates. So the switching rates are supposedly inherent individual traits that can be determined independently from total time T. Alternatively, the model can be transformed to use T square as the dependent variable by multiplying both hands of the model with T. Our efforts to establish a predictive model for final score failed. We found only a very weak correlation between scores and time. Also we weren't capable of demonstrating any relationships between the pre-test results and the final scores.

In view of the statement that there are "lies, damned lies and statistics", which is generally attributed to former British Prime Minister Benjamin Disraeli, we need to critically evaluate the significance of our findings. First of all this study is handicapped by having only log data and final scores available, while additional background profiles of students fail. Also, we had to do without tailored pre-tests and post-

tests, questionnaires, direct observations, and a randomised trial with experimental groups and a control group. Second, our starting point of associating the learners' increased freedom of movement with increased behavioural variability may be valid as such, but what is actually meant with freedom of movement and how it is expressed both in a qualitative and quantitative way is open to debate. Third, while we were able to identify switching behaviour as a relevant explanatory factor, we didn't provide an appropriate interpretation and theoretical foundation of this variable and a connection to cognitive states. Switching behaviour could be conceived as either a positive trait reflecting focused attention and efficient behaviours or a negative trait associated with impatience, superficiality or disorientation. As can be derived from Table 1 the average switching rates are typically 10 items per hour and are nothing like the frequent clicking that is common in social media and entertainment games. Fourth, the generalisation of the findings in this study is not straightforward because (naturally) all data are inherently tied to the specific game contents, game design and group of users. A different game play using different resources, videos, locations or any other game object will inevitably produce different effects and relationships.

A disclaimer putting these comments in perspective is that the purpose of our study was to explore to what extent the logging data of these particular games would be helpful to reveal meaningful patterns, variables and relationships. Even without the opportunity of collecting additional user information (e.g. profiles, appreciations, attitudes) or assigning different user groups to different conditions we have demonstrated the rich potential of game logging data and brought some relevant phenomena to the surface. Next steps in research call for developing and validating metrics for expressing game conditions such as freedom of movement, and use patterns such as switching behaviour, and make theoretical foundations for these in order to allow for empirical research and generalisation of findings. Exactly serious gaming could become an exemplary case of exploiting learning analytics.

References

1. VandeWalle, D., Brown, S.P., Cron, W.L., Slocum, L.W.: The influence of goal orientation and self-regulation tactics on sales performance: A longitudinal field test. Journal of Applied Psychology 84 (1999) 249–259.
2. Fisher, S.L., Ford, J.K.: Differential effects of learner effort and goal orientation on two learning outcomes. Personnel Psychology 51 (1998) 397–420.
3. Redeker, C., Punie, Y., Ferrari, A.: eAssessment for 21st century learning and skills. In: Ravenscroft, A., Lindsteadt, S. Kloos, C.D., Hernandez-Leo, D. (eds.). 21st Century Learning for 21st Century Skills. Proceedings of the 7th European Conference on technology-enhanced learning EC-TEL. Springer, Heidelberg (2012) 292-305.
4. Westera, W., Nadolski, R., Hummel, H., Wopereis, I.: Serious Games for Higher Education: a Framework for Reducing Design Complexity. Journal of Computer-Assisted Learning 24-5 (2008) 20-432.
5. Buckingham Shum, S., Ferguson, R.: Social Learning Analytics. Educational Technology & Society 15-3 (2012) 3–26.
6. Baker, R.S.J.D., Yacef, K.: The state of educational data mining in 2009: A review and future visions. Journal of Educational Data Mining 1-1 (2009) 3-17.

7. Bienkowski, M., Feng, M., Means, B.: Enhancing Teaching and Learning Through Educational Data Mining and Learning Analytics: An Issue Brief. SRI International U.S. Department of Education, Office of Educational Technology, Washington (2012). Retrieved from http://www.ed.gov/edblogs/technology/files/2012/03/edm-la-brief.pdf

8. Arnold, K.E.: Signals: Applying academic analytics. Educause Quarterly 33-1 (2010) 10-15. Retrieved from http://www.educause.edu/ero/article/signals-applying-academic-analytics

9. Pistilli, M., Arnold, K.: Course signals at Purdue: Using learning analytics to increase student success, 2nd International Conference on Learning Analytics and Knowledge. Vancouver (2012).

10. Long, P.D., Siemens, G.: Penetrating the Fog: Analytics in Learning and Education. Educause Review 46-5 (2011) 31-40. Retrieved from http://www.educause.edu/ero/article/penetrating-fog-analytics-learning-and-education

11. Boyd, D., Crawford, K.: Six Provocations for Big Data, A Decade in Internet Time: Symposium on the Dynamics of the Internet and Society, Oxford Internet Institute, Oxford (2012). Retrieved from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1926431

12. Gee, J.P.: What video games have to teach us about learning and literacy. Palgrave Macmillan, New York (2003).

13. Westera, W., Hommes, M.A., Houtmans, M., Kurvers, H.J.: Computer-Supported Training of Psychodiagnostic Skills. Interactive Learning Environments 11-3 (2003) 215-231.

14. Kickmeier-Rust, M.D., Albert, D.: Micro Adaptivity: Protecting Immersion in Didactically Adaptive Digital Educational Games. Journal of Computer Assisted Learning 26 (2010) 95-105.

15. Serrano-Laguna, A., Torrente, J., Moreno-Ger, P., Fernández-Manjón, B.: Tracing a Little for Big Improvements: Application of Learning Analytics and Videogames for Student Assessment, Procedia Computer Science 15 (2012) 203-209. Retrieved from http://www.sciencedirect.com/science/article/pii/S1877050912008344

16. Shute, V.J., Ventura, M., Bauer, M., Zapata-Rivera, D.: Melding the Power of Serious Games and Embedded Assessment to Monitor and Foster Learning: Flow and Grow. In: Ritterfeld, U., Cody, M., Vorderer, M. (eds.). Serious Games: Mechanisms and Effects. Routledge, New York (2009) 295-321.

17. Mislevy, R.J., Steinberg, L.S., Almond, R.G.: On the structure of educational assessment. Measurement: Interdisciplinary Research and Perspective 10 (2003) 3-62.

18. Pearl, J.: Probabilistic reasoning in intelligent systems: Networks of plausible inference. Kaufmann, San Mateo, CA (1988).

19. Reese, D.D., Seward, R.J., Tabachnick, B.G., Hitt, B., Harrison, A., McFarland, L.: Timed Report Measures Learning: Game-Based Embedded Assessment. In: Ifenthaler, D., Eseryel, D., Ge, X. (eds.). Assessment in Game-Based Learning: Foundations, Innovations, and Perspectives. Springer, New York (2013).

20. Green, S.B.: How many subjects does it take to do a regression analysis? Multivariate Behavioral Research 26 (1991) 499–510.