

Using Game Learning Analytics for Validating the Design of a Learning Game for Adults with Intellectual Disabilities

Ana R. Cano, Complutense University of Madrid. Calle del Prof. José G^a Santesmases, 9, 28040 Madrid, Spain (Despacho 414) | (+34) 91 394 76 23 | Fax (+34) 91 394 75 29. anarcano@ucm.es

Baltasar Fernández-Manjón, Complutense University of Madrid. balta@fdi.ucm.es

Álvaro J. García-Tejedor, Francisco de Vitoria University. a.gtejedor@ceiec.es

Ana Rus Cano is a researcher at e-UCM group, Complutense University of Madrid, Spain. Her research interests are in the area of Serious Games for people with Intellectual Disabilities, Game Learning Analytics and e-Learning. Álvaro J. García-Tejedor is an AI professor at Francisco de Vitoria University and head of CEIEC, a research center for social application of information technologies. Baltasar Fernández-Manjón is a professor in CS at Complutense University of Madrid and director of the e-UCM research group and of the Telefonica-Complutense Chair on Serious Games. His research interests are in e-learning, application of Serious Games and Game Learning Analytics. Address for correspondence: Ms. Ana Rus Cano, Complutense University of Madrid, Computer Science Faculty, Prof. José García Santesmases 9, 28040 Madrid, Spain. Email: anarcano@ucm.es

Abstract

Serious Games, defined as a game in which education (in its various forms) is the primary goal rather than entertainment, have been proven as an effective educational tool for engaging and motivating students (Michael & Chen, 2006). However, more research is needed to sustain the suitability of these games to train users with cognitive impairments. This empirical study addresses the use of a Serious Game for training students with Intellectual Disabilities in traveling around the subway as a complement to traditional training. Fifty-one (51) adult people with Down Syndrome, mild cognitive disability or certain types of Autism Spectrum Disorder, all conditions classified as intellectual disabilities, played the learning game *Downtown, A Subway Adventure* which was designed ad-hoc considering their needs and cognitive skills. We used standards-based Game Learning Analytics techniques (i.e. Experience API –xAPI), to collect and analyze learning data both off-line and in near-real time while the users were playing the videogame. This article analyzes and assesses the evidence data collected using analytics during the game sessions, like time completing tasks, inactivity times or the number of correct/incorrect stations while traveling. Based on a multiple baseline design, the results validated both the game design and the tasks and activities proposed in *Downtown* as a supplementary tool to train skills in transportation. Differences between High-Functioning and Medium-Functioning users were found and explained in this paper, but the fact that almost all of the students completed at least one route without mistakes, the general improvement through sessions and the low-mistake ratio are good indicators about the appropriateness of the game design.

Keywords

serious games, learning games, game learning analytics, intellectual disabilities, down syndrome

Introduction

The use of technological tools for educational purposes is becoming popular when training people with Intellectual Disability (ID) in daily routines. Serious Games are an example of these emerging tools. However, more research is needed to sustain the suitability of these games to train users with cognitive impairments. Recent researches explored the effectiveness of the Serious Games as learning mechanisms for users with intellectual disorders, obtaining positive outcomes in their investigations (Kwon & Lee, 2016) (Cano, Fernández-Manjón, & García-Tejedor, 2015) (Chang, Kang, & Liu, 2014) but most researches imply a very limited number of users and usually only feature in qualitative studies.

ID is the impairment characterized by significant limitations in both intellectual functioning and in adaptive behavior, which covers many everyday social and practical skills. Learning, problem solving, reasoning and activities of daily living (like personal care, obeying rules and laws or travel/transportation) are challenging skills for ID individuals (AAIDD American Association on Intellectual and Developmental Disabilities, 2010). Their delays in motor milestone attainment, sensorimotor performance deficit and perceptual dysfunctions, in addition to significant limitations both in intellectual functioning and in adaptive behavior are barriers that these individuals must face when learning complex tasks (Burack, Hodapp, & Zigler, 1988).

Although the use of Serious Games is increasing in educational environments, their proven effectiveness is still scarce. There is a lack of empirical research and methodologies that provide evidence of what types of developments are effective for educational purposes (Connolly, Boyle, MacArthur, Hainey, & Boyle, 2012). This fact is even more noticeable talking about games for students with ID for two reasons: 1) developing games for learners with cognitive disabilities are expensive and laborious due to the accessibility features that developers have to include in the game

design and 2) the population of students with ID is smaller compared to the number of learners available in general learning studies, therefore, researches involving individuals with ID are limited to smaller pilots and user evaluations (Roozeboom, Visschedijk, & Oprins, 2017).

The purpose of this research is to provide insight about the use of Game Learning Analytics (GLA) as a measure mechanism to validate the effectiveness of a Serious Games designed specifically for learners with Intellectual Disabilities.

The results were obtained from an experiment involving 51 adult people with ID (like Down Syndrome, mild cognitive disability or certain types of Autism Spectrum Disorder -ASD) that played the learning game *Downtown, A Subway Adventure*. The game design principles used in *Downtown* are explained in the first section. Then, we describe how we used Game Learning Analytics for gathering data from the game. Next, the methodology of the experiment is also included in this paper as an example of best practices while developing and testing an accessible learning game for users with ID. Data analysis obtained and conclusions are available at the end of the article.

Downtown, A subway Adventure

Downtown: A Subway Adventure is a spy game developed for students with ID, between 18 and 45 years old. The game is designed to train individuals in using the public subway transportation system. *Downtown* simulates the subway of Madrid (Spain) in a 3D realistic perspective (see figure 1). The aim is to help game players to identify the game environment with reality when they are travelling by their own (thus facilitating the transfer knowledge).



Figure 1: Detail of the train wagon (left) and subway station (right) in *Downtown*

An adequate game design is crucial for the development of the research. *Downtown* is a game especially designed for players with ID, considering their intellectual, psychological and motor characteristics that impact in their learning abilities. Despite the fact that the cognitive skills can vary from one individual to another, there are common characteristics that provide guidance and rules about the needs of the users while playing.

The basic design principles we followed for the game design are based on best practices, literature review, developers' previous experience and experts' advice (psychologists and trainers specialized in ID) that participated in the initial game design. There are five aspects that we included in *Downtown's* design: 1) the game should be realistic, based on real life. The virtual environment had to be as similar as possible to the real one so they can easily recognize and translate what is happening in the game to the real environment, 2) the game accessibility guidelines (Ablegamers Foundation, 2012-2017) have to be implemented in the game at different levels (advanced degree for cognitive and general dimensions and intermediate degree for motor, vision, hearing and speech). We followed the rules provided by the Game Accessibility Guidelines, that brings together a group of design decisions that impact in the mechanics of the game (e.g. number of options available for the player at a given time), 3) educators, trainers and psychologist specialized in ID should be involved from the very beginning of the game concept, 4) the mechanics of the game should be accommodated to the learners' impairments and 5) PC and mouse should be the platform and input device to play, because these are the tools that they currently use in the technology classes.

The mechanics of the game were specifically designed taking into account the suggestions of the experts involved in the actual training in the subway to address the common problems and situations that ID users face when they are using the subway: what to do if they take a wrong path, how to act if a stranger talks to them, what to do if they fall asleep, etc. *Downtown* not only trains the students for choosing the right route while traveling from one metro station to another, but also includes puzzles and missions that help them improve basic daily skills like independence, long and short memory or spatial vision. Some of the social aspects that are complex for them to master are also included (e.g. interact with the subway operators if the transport pass does not work). Further information about the game design is available in a previous publication (Cano, Fernández-Manjón, & García-Tejedor, 2016).

The use of Game Learning Analytics for Inclusive Serious Games

The approach for developing *Downtown* was a user-centered design, where both ID students and trainers from Down Madrid were actively involved in the design of the game mechanics and interfaces. There were four testing sessions before releasing the final fully-playable game. Although the opinions of users were key to assuring the adaptability of the game, ID people deal with strong communication problems so traditional game evaluation methods (e.g. pre-post test, user questionnaires) could not be always fully reliable. Serious Games can be used as a measurement learning tool (Kato & de Klerk, 2017), giving trainers and instructors the opportunity of gathering data about the user's play patterns within the game. When using an educational videogame, it is possible to track the evolution and engagement of the students and use that data to better understand, or even predict, the learning outcomes. The application of Learning Analytics is defined as "the measurement, collection, analysis and report of data about learners and their context, for purposes of understanding and optimizing learning and the environments in which it occurs" (Long & Siemenes, 2011). The application of LA in educational videogames is called Game Learning Analytics (GLA) and combines the educational goals of LA with the tools and technologies from Game Analytics (Freire, et al., 2016). Gathering data about the learning process of ID users directly from the videogame can be a powerful tool to validate a Serious Game compared to observational methods. ID users struggle with several communication issues that can be a barrier for educators while trying to understand if they are processing the information presented in the game and if they are having an effective learning experience within the game.

Data Tracker

Data traces were collected using a tracker included in the game. The tracker is used in the game to send out the relevant information in near-real time about the behavior of the users and their learning patterns while playing the game. The tracker used is open source and has been developed by the H2020 RAGE project (RAGE H2020 Project, 2014-2017). All the data gathered follows the standard xAPI (Experience Application Programming Interface) a new emerging specification for collecting, storing and reporting user interactions on learning systems (Serrano-Laguna, et al., 2017). While the user is playing, the tracker is used to send out traces to an online server that performs different analysis based on what different types of views in a near real-time web-browser dashboard are built. The server provides visualizations about the performance of the classroom that is playing at that moment, or individualized reports for each user. Trainers and researchers were monitoring the dashboards during the sessions, to provide help and feedback to the users while are playing *Downtown*. As we mentioned before, ID individuals struggle with communication so following their activity inside the game through an indirect source, like the server's dashboards, provide researchers valuable information about the user's learning performance supplementary to the direct observation. The game analytics gathers two types of observables: 1) data related to the game options and 2) data related to the user interaction with the game. The first type allows the researchers to know the needs of the users related to the interface, depending on their disability features, and the second one provides the information about their learning performance while using the game.

Methodology

Participants

In order to validate the design of the game *Downtown*, we run a case study in cooperation with Fundación Síndrome de Down de Madrid (Down Madrid) including fifty-one (51) adults, ages between 19 and 41, with diverse types of Intellectual Disabilities. The case study consists of playing the game for 3 one-hour sessions. Data from nine users were discarded because, for different reasons, they didn't complete all the game sessions ($n=42$, $M_{age}=29$, $SD_{age}=7.07$). All students were previously enrolled in the introduction to technology classes organized by Down Madrid. Because the game is designed to High and Medium-Functioning Down Syndrome and Mild Intellectual Disability (MID) individuals, as an inclusion criteria users were screened according to their degree of autonomy, self-confidence and ability to accomplish daily tasks. Low-functioning users were discarded. From the participants included in the study, 71% were Down while 29% had other disabilities, like MID or different types of Autistic Spectrum Disorder (some Down participants can present other co-occurrent cognitive impairments). All sessions were videotaped just in case some data-driven results need to be reviewed or contrasted with the actual class experience.

Randomization and Population's Cognitive Abilities

There were six different technology classes in Down Madrid where users were divided randomly, depending on their schedule availability. Individuals in each group have different IQ, cognitive competences and autonomy. Before the sessions took place, trainers were asked to complete information about each student. The data provided helped researchers understand the learning capabilities of each individual and are used as an input for integrating and merging data from the qualitative survey and the runtime results from game sessions. Information collected about each student covers three aspects that can influence in the learning process while playing the game: 1) neurocognitive functions, based on six domains defined by the SSA, the United States Social Security Administration, 2) previous experience using city-public transportation and 3) expertise in the use of technology and videogames.

The six cognitive domains relevant to SSA are: general cognitive/intellectual ability, language and communication, memory acquisition, attention and distractibility, processing speed, and executive functioning (OIDAP, Occupational

Information Development Advisory Panel, 2009). Each aspect was surveyed with a 5-point Likert scale (1 = *very low*, 2 = *low*, 3 = *medium*, 4 = *high*, 5 = *very high*). Reliability of the scale was good according to Cronbach's α coefficient (11 items, $\alpha = .93, >.8$) which means that the reliability of the method is considered adequate.

Based on the test completed by the trainers, we can differentiate between two types of users: Medium-Functioning users (SSA avg. score ≤ 3) and High-Functioning users (SSA avg. score > 3). This classification is based on the current practice for grouping the users in the technology classes and can be useful to identify differences in how the users are playing depending on their cognitive features.

Trainers also completed information about the age of the students and their physical appearance to analyze whether they choose a videogame character similar to themselves and if that can promote learning transfer from the game to reality (Klimmt, Hefner, Vorderer, Roth, & Blake, 2010). Full explanations and references about this topic are available in the *Data Analysis* section, further in this article.

Sessions

Fifty-one participants (51) played the game *Downtown, A Subway Adventure* in the Down Madrid facilities during May and June 2017. Sessions length were one hour. As previously stated, students were asked to play the game for at least three hours to be finally included in the case study.

ID individuals present difficulties in attentional control and literacy that impact in the use of the game and the learning process (Lanfranchi, Cornoldi, & Vianello, 2004). To minimize the effect of these conditions, two researchers and two trainers were helping the users during the sessions (see figure 2).



Figure 2: Trainers and researchers helping the students during a game session

Participants began the sessions by listening to instructions about the operational mode of the game: controls, accessibility options, interface and an explanation about the tasks that they must accomplish. Each game session was individual: researchers provided an anonymous token to each user. Users have the same token in all of the game sessions that they play. The game sessions are identified by tracking both the timestamp when the game starts and the token of the user.

Then, they are asked to create and customize the main character of the videogame. They can choose between different types of hair styles, faces, bodies and apparel according to their personal preferences (see figure 3).



Figure 3: Down character selected in customization screen

Students must navigate the accessibility menu provided by the game before starting to play with supervision from the trainers. They are asked to modify the pace of the game, the text size, the colors of the text and background and the volume of the dialogs and music. By adapting the accessibility options specifically for each individual, researchers make sure that the interface of the game is not interfering in the gameplay and the knowledge acquisition.

Once they are comfortable with the game environment, users start playing the game. The game offers four different levels (easy, medium, hard and expert). All users start playing in the easy level, and progress to the higher one once the previous is completed. Origin and destination subway stations, number of transfers and complexity of the route is assigned automatically by the game depending on the difficulty level chosen by the player/trainer.

We consider that one level is completed when all the tasks, missions and minigames included in the level are completed.

Downtown is a replica of Madrid (Spain) subway network where players can travel around the stations as they wish simulating the real environment. At the beginning of the game, users start their routes in a random station and are asked to travel to another one. Like in real life, they need to plan the route using the metro maps, choose the transfer stations and wait for the trains.

Observables

Data collected was anonymized and only linked with each user by the token provided at the beginning of the sessions. Only Down Madrid trainers (but not researchers) can relate each user with the game data.

The game doesn't collect any personal information. The variables collected about the user interaction are: character preferences, accessibility preferences, user id data, game session total time, time completing a session/minigame, total inactivity time, attempts to complete a minigame (number of fails before success), time traveling from origin to destination, number of correct/incorrect stations during the route, number of clicks in the accessibility menu during the game, number of clicks in "help" items, time completing tasks after help, number of attempts completing tasks after help and percentage of game progress for every timestamp.

Data analysis

The case study provides insight about the effectiveness of *Downtown* using an evidence-based model based on the data collected. More than three hundred thousand traces (i.e. interaction data) were gathered and analyzed using Game Learning Analytics techniques. To organize and guide the analysis, we proposed six questions each with an associate hypothesis. The questions were determined by an examination of the literature available as a result of observational methods (not data-based driven):

Q1: Is the identification with the avatar an important issue when learning with an educational game?

Q2: Do the cognitive skills of the users impact in their performance with the game?

Q3: Do users with previous transportation training obtain better results in the game than users without training?

Q4: Do heavy-technology users learn faster how to play with the videogame?

Q5: Are videogames a motivational learning environment for users with ID?

Q6: Does the game design impact on the way the users interact with the videogame?

To address these questions, the following hypothesis are tested:

H1: Users that identify themselves with the avatar obtain better game results than users choosing a random character

H2: High-Functioning users do a better performance using the game than Medium-Functioning users

H3: Users with previous experience in transportation training have a better performance using the game

H4: Users that play videogames on a regular basis record better performances using the game

H5: ID users are engaged and motivate while learning with a videogame

H6: *Downtown's* game design is effective as a learning tool

H1: Users that identify themselves with the avatar obtain better game results than users choosing a random character

Literature suggests that the more similar a videogame character is to each player, the easiest the learning transfer from the game to reality is and the best results in learning games they obtain (Klimmt, Hefner, Vorderer, Roth, & Blake, 2010) (Newman, 2002) (Griebel, 2006).

Data analyzed from *Downtown* showed that most of the users selected the preconfigured character despite the fact that they were asked to customize the avatar at the beginning of each game session. Thirty users (71.4%) stayed attached to their avatars while 12 (28.5%) users changed their avatars from one session to another. Trainers and researchers encouraged them to modify the main character and assisted the users in the process. None of the users selected the avatar with Down features, despite the trainers showed them the avatar and pointed that the avatar was Down.

After the analysis, we did not observe significant evidence in the user's play patterns between those players who customized the character and those who didn't: average time completing levels, inactivity times and success/mistakes

ratio are alike between the two groups of users. Our evidence-based analytics data initially refutes the statement presented in previous research (that mainly was derived from external observation) suggesting that more research is needed on the topic.

H2: High-Functioning users do a better performance using the game than Medium-Functioning users

As previously stated, cognitive abilities, skills and autonomy of the users were obtained using the SSA scale with the help of the trainers. Users were divided in two groups: 19 students (45.2%) were considered Medium-Functioning (MF) users and 23 (54.8%) were considered High-Functioning (HF) users depending on their SSA scores. Both groups of users are comparable in number for the analysis.

Data shows that 86.9% of the HF users were able to complete the easiest level of the game and played the medium level compared to 36.8% of the MF users (We used the chi-square statistic to compare our dependent variables: $\chi^2 = 11.3815$ $p\text{-value} = 0.00074 < 0.05$. The results are statistically significant). Only 5.2% of the MF finished the medium level versus 39.13% of the HF users ($\chi^2 = 6.5787$ $p\text{-value} = 0.01032 < 0.05$. The results are statistically significant). However, despite the first impression of the trainers, average times completing tasks for MF users is almost the same compared to HF users while playing the easiest level (See Figure 4). Note that as the cognitive abilities can highly vary from one user to another, the game was designed to offer very simple and intuitive tasks in the easiest levels.

Differences appear in medium level, where MF are not able to finish the tasks proposed or take too long to complete them. Some of the blockages they encounter are: difficulty finding the right paths, confusion when finding and reading signals or disorientation when changing lines in the same station. This fact may occur for two different reasons: there are deficiencies in the game design (tasks are too complex) and/or the learning curve for MF users is slower compared to HF users, which means that they need more time to complete the tasks. According to the trainers' opinion, a difference of almost 7 minutes of gameplay is relatively high. Note that tasks in medium level are designed to be accessible enough for both groups, so there shouldn't be differences between MF and HF performances. Neither of the only two MF users that completed the medium level could finish the levels hard or expert.

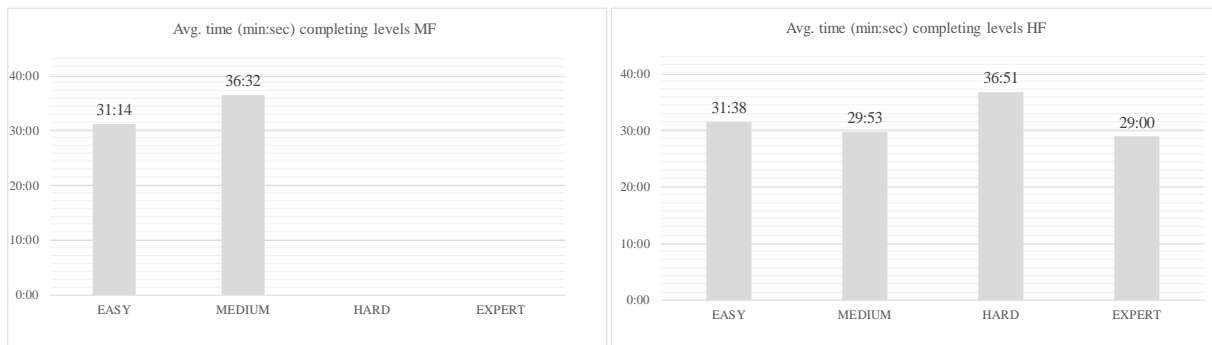


Figure 4: Average time (min:sec) completing levels for Medium-Functioning users vs. High-Functioning users

H3: Users with previous experience in transportation training have a better performance using the game

Downtown was developed as a supporting learning tool to include as supplementary content in the traditional transportation training provided by Down Madrid. Some of the users that played the game participated in previous transportation trainings and already knew how to move around the city using the public transportation system (17.6% of the total game population).

After playing *Downtown*, no significative differences were observed between previously trained and untrained students while playing the game regarding time completing tasks, inactivity times and success/mistakes ratio. Data shows that the ability demonstrated when using the videogame or completing the tasks proposed by the game is independent of the previous experience using public transportation.

H4: Users that play videogames on a regular basis record better performances using the game

40.5% of the students included in the research play videogames at home. Trainers considered them heavy-technology users based on their previous experience in class. They are used to handle mobile devices, PCs and consoles as neurotypical users. In this instance, data confirm that users that play commercial videogames in a regular basis completed the levels faster than non-player users. Average time completing tasks for videogame players is 28:18 minutes which represent a difference of 12% less time than non-players (see figure 5).

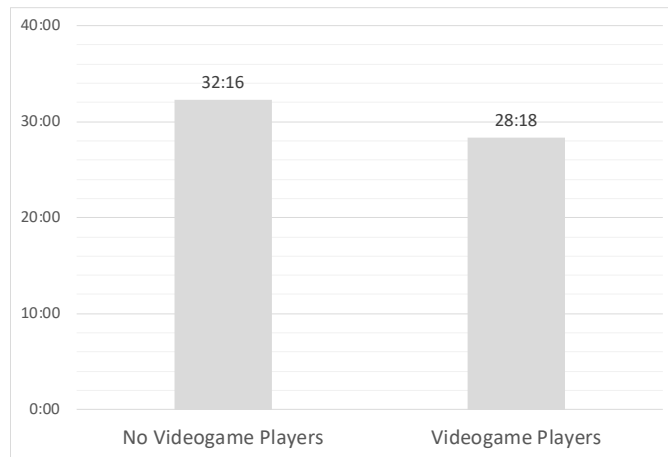


Figure 5: Average time completing tasks (min:sec): videogame players vs. non-players

There is also evidence that students that play videogames have less inactivity times and made less mistakes than non-players.

H5: ID users are engaged and motivated while learning with a videogame

Motivation and engagement are crucial in the learning acquisition process. ID users are slow learners due to their difficulties in the process of abstractions, conceptualization, generalization and learning transfer (Shallock et al., 2010). According to their trainers, they experience a lack of motivation in a short period of time, especially if the tasks to complete are difficult or monotonous.

We evaluated the user engagement analyzing the average inactivity time during game sessions (objective, using the data tracker) and reviewing the videos of the game sessions (subjective). We define inactivity time as the time elapsed between two events inside the game. The events in *Downtown* are automatic and related to tasks, which means that as soon as one task is completed another event pops-up. All the events are designed to be completed in a certain period of time, so we assume that if a new event is not appearing after that time, the user is “inactive” in the game. Inactive periods of time may be an indicator of discouragement.

We also reviewed the videos of the sessions, matching the heavy inactivity times with the records. Some users used the bathroom while playing the videogame or asked the trainers for help. That data was excluded from the analysis.

Data shows that the average inactivity time is reduced by a 57.7% percentage average from session 1 to session 8 (see figure 6), which means that the users improved in the use of the videogame from one session to the following and, in combination with the decrease of mistakes through sessions, may suggest that the more they play the more engaged in the game they are.

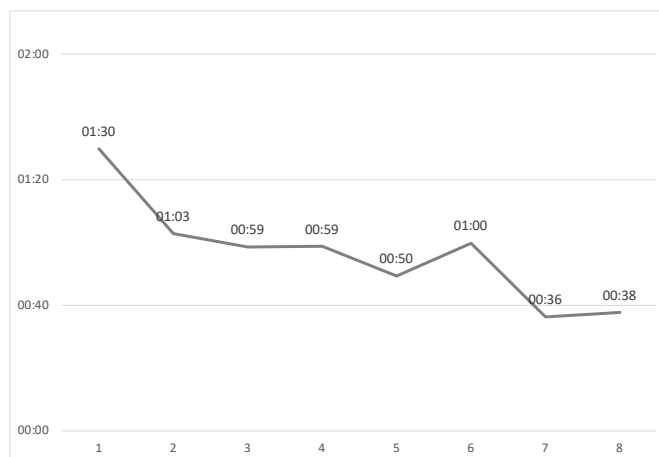


Figure 6: Average inactivity time through sessions (min:sec)

H6: Downtown’s game design is effective as a learning tool

Latest research agree with the statement that Serious Games provide more learning benefits compared to traditional teaching approaches, but empirical and data-based studies proving the effectiveness of serious gaming are still scarce (Roozeboom, Visschedijk, & Oprins, 2017) (Torrente, del Blanco, Moreno-Ger, & Fernández-Manjón, 2012). We

consider *Downtown* a supplementary content to the traditional training, not as a learning object per se, but our data can provide information about the suitability of the game design's decisions made. Designing serious games for ID students is complex and laborious because every mechanic must be adapted to their cognitive features and the intellectual skills of each individual are different.

Our data reveal promising results about how ID users can learn to interact with a serious game. All the trainers agree that the use of *Downtown* would enhance the users' learning acquisition (Perceived Usefulness) (Davis, 1989) and plan to include it as supplementary content in the traditional training.

Each game sessions lasted one hour and during that time, students were asked to complete as many routes as they could. Each route takes between 10 (easy level) and 40 (expert level) minutes to complete. The higher the level, the longer the path and, consequently, the more time spent playing in that level. We decided to adjust the time for each level depending on two factors: 1) the advice of the psychologists and trainers and 2) the time that it would take to complete the route in real life.

During the sessions, all of the students were allowed to rest, go to the bathroom, talk with the trainers, etc. to minimize the effect of mental fatigue. Thus, the number of routes completed by each individual was different. Most of the users (85.8%) were able to reach a destination, by following the right path, proposed by the game (both MF and HF).

Half of the mistakes and wrong paths (50.8%) occurred during the first 30 minutes of effective playing, once they completed 2 or 3 routes (see figure 7).

Since the difficulty of the paths proposed increase in each level, the fact that the mistakes are made at the beginning of the sessions may suggest that the users are learning how to interact with the game.

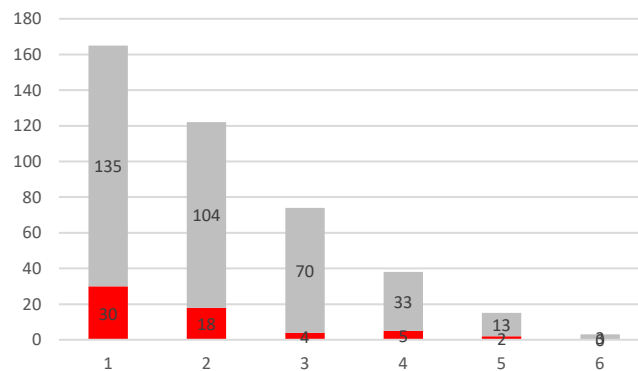


Figure 7: number of correct (grey) vs incorrect stations (red) per game session

Once that they know how the video game works the routes are more accurate and the mistakes rate decrease. These figures suggest that users are improving in the use of the game.

The fact that almost all the students completed at least one route without mistakes means that the instructions and tasks proposed in the game are clear and well-adapted to a wide range of ID users, despite their cognitive abilities. Moreover, all the students improved their performance from one session to another, although the complexity of the tasks increase in higher levels.

Conclusion

This paper describes the methodology and main outcomes of the case study to validate the game design of *Downtown* using Game Learning Analytics techniques. Data obtained allow us to contrast the design decisions to validate the game. After reviewing the data, the results validated both the game design and the tasks and activities proposed in *Downtown* as a potential supplementary tool to train skills in transportation. There is evidence based on the analytics data that the game is adapted to the cognitive skills of the users despite the very variable degree of their intellectual abilities.

Both Medium-Functioning and High-Functioning individuals did experience improvement performing the activities proposed by the game despite the negative initial impression of the trainers, who thought that the only ones who would improve were the HF users. All the trainers agree that the use of *Downtown* would enhance the user learning acquisition.

Although the results show a tendency for users to use the same character in all the sessions, we didn't observed differences in the play patterns between the players that customized the character and those who didn't. Future research will have to show if hypothesis number one is valid with users that play the game before doing the actual training in the subway.

HF users were able to complete more levels in a shorter period of time than MF users. HF users perform the tasks proposed by *Downtown* faster and with less errors than MF users. These facts are consistent with the number of mistakes and the average inactivity times of HF users, that are lower.

Results suggest that users need to understand how the videogame works (mechanics, interface, controls) before starting to assimilate the game content. Previous training on transportation seems to have no impact on game performance. Users that were previously trained commented things like “the subway looks so real” or “I travel this line every day” but their performances inside the game were not different compared to untrained users. As expected, users that play videogames completed the levels faster than non-players.

Trainers consider the game a positive and motivational learning environment: almost all the users show improvement and engagement performing the videogame tasks. The fact that almost all the students completed at least one route without mistakes, the gameplay improvement through sessions and the low-mistake ratio are good indicators about the appropriateness of the game design.

In future studies, we would like to confirm the value of the game as a learning tool based on learning outcomes after playing the game. Phase two will take place in Madrid during December 2017. A group of users that played the game in phase one will combine on-site training in the subway with game sessions and will be compared to untrained players. Data will be collected, analyzed and compared to pre-test and post-test data collected by the trainers. The results of this further investigation will give more information about the learning patterns, outcomes and the effectiveness of the game as a learning tool.

Statements on Open Data, Ethics and Conflicts of Interest

Full research protocol was examined by the Down Madrid Ethics Committee before the beginning of the study and approval was granted. Down Madrid included the testing of *Downtown* in their mobility program. All of the ID students and their families were carefully informed and agree to participate in the study while enrolling in the technology classes. Furthermore, all the data gathered were anonymized and exclude any personal information (besides the age of the users, for statistical purposes). Only Down Madrid, not researchers, can match the information analyzed with each student, as described in section “Observables”, in order to help them in the training process. Informed Consent between researchers and Down Madrid is attached when submitting this paper.

Rights and exploitation over data belong exclusively to Down Madrid and Complutense University of Madrid and can be only used for research purposes defined in the experimental design. Data is stored in the e-UCM group repository and is available upon request. For more information contact: anarcano@ucm.es

The authors declare no conflict of interest.

Acknowledgements

The authors thank the assistance of Maria Luisa Berdud, Director of the TIC Service in Fundación Síndrome de Down de Madrid, for her assessment in choosing the users, providing the facilities for the research and her commitment with the project. The CEIEC Institute has been partially funded by the Francisco de Vitoria University (UFV-2015-05). This work has been partially funded by Regional Government of Madrid (eMadrid S2013/ICE-2715), by the Ministry of Education (TIN2013-46149-C2-1-R, TIN2017-89238-R) and by the European Commission (RAGE H2020-ICT-2014-1-644187, BEACONING H2020-ICT-2015-687676, Erasmus+ IMPRESS 2017-1-NL01-KA203-035259).

References

- AAIDD American Association on Intellectual and Developmental Disabilities. (2010). *Intellectual Disability: Definition, Classification, and Systems of Supports (11th Edition)*. doi:ISBN 978-1-935304-04-3
- Ablegamers Foundation. (2012-2017). *Game Accessibility Guidelines*. Retrieved March 09, 2016, from <http://gameaccessibilityguidelines.com/>
- Burack, J., Hodapp, R., & Zigler, E. (1988). Issues in the classification of mental retardation: Differentiating among organic etiologies. *Journal of Child Psychology and Psychiatry*, 29, 765-779.
- Cano, A., Fernández-Manjón, B., & García-Tejedor, Á. (2015). Highlights in the Literature Available in Serious Games for Intellectual Disabilities. *ICWL International Conference on Web-Based Learning*. Guangzhou, China.
- Cano, A., Fernández-Manjón, B., & García-Tejedor, Á. (2016). *Downtown, A Subway Adventure: Using Learning Analytics to Improve the Development of a Learning Game for People with Intellectual Disabilities*. *ICALT Advanced Learning Technologies*. Austin, Texas: IEEE Computer Society Publications.
- Chang, Y., Kang, Y., & Liu, F. (2014). A computer-based interactive game to train persons with cognitive impairments to perform recycling tasks independently. *Research in Developmental Disabilities*, 35(12), 3672-3677. Retrieved from <https://doi.org/10.1016/j.ridd.2014.09.009>
- Connolly, T., Boyle, E., MacArthur, E., Hainey, T., & Boyle, J. (2012, September). A systematic literature review of empirical evidence on computer games and serious games. *Computers & Education*, 59(2), 661-685.
- Davis, F. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly*, 13(3), 319-340. doi:10.2307/249008
- Freire, M., Serrano-Laguna, A., Manero, B., Martínez-Ortiz, I., Moreno-Ger, P., & Fernández-Manjón, B. (2016). Game Learning Analytics: Learning Analytics for Serious Games. *Learning, Design, and Technology*, 1-29.

- Griebel, T. (2006). Self-Portrayal in a Simulated Life: Projecting Personality and Values in The Sims 2. *The international journal of computer game research*, 6(1). Retrieved from <http://www.gamestudies.org/0601/articles/griebel>
- Kato, P., & de Klerk, S. (2017). Serious Games for Assessment: Welcome to the Jungle. *Journal of Applied Testing Technology*, 18 (S1), 1-6.
- Klimmt, C., Hefner, D., Vorderer, P., Roth, C., & Blake, C. (2010). Identification With Video Game Characters as Automatic Shift of Self-Perceptions. *Media Psychology*, 13(4), 323-338.
- Kwon, J., & Lee, Y. (2016). Serious games for the job training of persons with developmental disabilities. *Computers & Education*, 95, 328-339. Retrieved from <https://doi.org/10.1016/j.compedu.2016.02.001>
- Lanfranchi, S., Cornoldi, C., & Vianello, R. (2004, Nov). Verbal and visuospatial working memory deficits in children with Down syndrome. *American Journal of Mental Retardation*, 109(6), 456-466. doi:10.1352/0895-8017(2004)109<456:VAVWMD>2.0.CO;2
- Long, P., & Siemenes, G. (2011). Penetrating the Fog: Analytics in Learning and Education. *1st International Conference on Learning Analytics and Knowledge*. Banff, Alberta, Canada.
- Michael, D., & Chen, S. (2006). *Serious Games: Games That Educate, Train, and Inform*. Boston, Mass. : Thomson Course Technology, c2006.
- Newman, J. (2002). The Myth of the Ergodic Videogame. *The international journal of computer game research*, 2(1). Retrieved from <http://gamestudies.org/0102/newman/>
- OIDAP, Occupational Information Development Advisory Panel. (2009, June 8). *Mental cognitive subcommittee: Content model and classification recommendations*. Retrieved from <https://www.ssa.gov/oidap/Documents/AppendixC.pdf>
- RAGE H2020 Project. (2014-2017). *RAGE Realising an Applied Gaming Ecosystem*. Retrieved from <http://rageproject.eu/>
- Roozeboom, M., Visschedijk, G., & Oprins, E. (2017, January). The effectiveness of three serious games measuring generic learning features. *British Journal of Educational Technology*, 48(1), 83-100. doi:10.1111/bjet.12342
- Serrano-Laguna, A., Martínez-Ortiz, I., Haag, J., Regan, D., Johnson, A., & Fernández-Manjón, B. (2017). Applying standards to systematize learning analytics in serious games. *Computer Standards & Interfaces* 50, 116-123. Retrieved from <http://dx.doi.org/10.1016/j.csi.2016.09.014>
- Shalock et al. (2010). *Intellectual Disability: Definition, Classification and Systems of Supports. 11th Edition*. AAIDD (American Association on Intellectual and Development Disabilities).
- Torrente, J., del Blanco, A., Moreno-Ger, P., & Fernández-Manjón, B. (2012). Designing Serious Games for Adult Students with Cognitive Disabilities. *19th International Conference on Neural Information Processing (ICONIP2012)*. Doha.