

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2019.Doi Number

Game Analytics Evidence-Based Evaluation of a Learning Game for Intellectual Disabled Users

Ana Rus Cano¹, Álvaro García-Tejedor², Cristina Alonso-Fernández¹ and Baltasar Fernández-Manjón¹, Senior Member, IEEE

¹ Department of Software Engineering and Artificial Intelligence, Complutense University of Madrid, Spain

² CEIEC Centro de Estudios e Innovación en Gestión del Conocimiento, Francisco de Vitoria University, Madrid, Spain

Corresponding author: Ana Rus Cano (e-mail: anarcano@ucm.es).

ABSTRACT Learning games are becoming popular among teachers as educational tools. However, despite all the game development quality processes (e.g., beta testing), there is no total assurance about the game design appropriateness to the students' cognitive skills until the games are used in the classroom. Furthermore, games designed specifically for Intellectual Disabled (ID) users are even harder to evaluate because of the communication issues that this type of players have. ID users' feedback about their learning experience is complex to obtain and not always fully reliable. To address this problem, we use an evidence-based approach for evaluating the game design of *Downtown, A Subway Adventure*, a game created to improve independent living in users with ID. In this paper we exemplify the whole process of applying Game Analytics techniques to gather actual users' gameplay interaction data in real settings for evaluating the design. Following this process, researchers were able to validate different game aspects (e.g., mechanics) and could also identify game flaws that may be difficult to detect using formative evaluation or other observational-based methods. Results showed that the proposed evidence-based approach using Game Analytics information is an effective way to evaluate both the game design and the implementation, especially in situations where other types of evaluations that require users' involvement are limited.

INDEX TERMS evidence-based learning, game analytics, game design, game evaluation, intellectual disability, learning games, serious games

I. INTRODUCTION

Recently, Learning Games (also known as serious games, educational games or applied games) became popular learning tools so there are many research and application projects about this type of games [1]. Most of the literature available praises their advantages compared to traditional teaching methods, like the positive attitude of the students toward the games or the authentic highly interactive learning environment that they promote [2][3][4][5]. However, more research about how to create effective designs that optimize both the development process and the adequacy of the game mechanics to the users' cognitive abilities is still scarce [6] [7].

This problem is particularly evident when developing an educational game for people with Intellectual Disabilities (ID) [8]. The American Psychiatric Association defines intellectual disabilities as neurodevelopmental disorders that begin in childhood and are characterized by intellectual difficulties as well as difficulties in conceptual, social, and practical areas of living [9]. In particular, users with ID

present certain cognitive characteristics that affect the way they learn like limited memory, difficulty sustaining attention during long periods of time or confusion in the process of abstraction, conceptualization and transferring the conceptual learning to real settings [10].

ID students acquire skills and knowledge at a different pace than neurotypical learners and face a set of learning challenges that should be taken into account when designing learning games [11]. As ID users typically struggle with communication problems [12], the application of traditional observational methods for evaluating learning games is complex, expensive and even not fully reliable [13]. Serious game designing process is a complex task where there are not standardized methodologies to guarantee that the designers' ideas are adequately translated into playful game mechanics (e.g. missions, minigames) [14]. Beta testing and early user involvement are common formative evaluation practices to test and improve serious games [15], but some users, such as students with ID, have

specific peculiarities, like communication problems, that make these practices even more challenging.

As budget is usually very limited, and the game testing in ecological situations (real users, actual environment) is so complex and expensive, developers usually have no assurance about the effectiveness of their game designs and the adequacy of the final product to the users' skills. This often means that game shortcomings are identified too late after testing the final game in the actual class [16].

In this paper we present an evidence-based approach for evaluating the game design of *Downtown, A Subway Adventure*, developed for players with ID. The purpose of this game is to help the users in navigating in the complex Metro network of Madrid (a public subway system) and acquire skills to solve problems that can appear when traveling independently. We consider that this approach can be used as a reference and somewhat can be generalized to other game designs and developments created for users with ID. The game includes the use of Game Analytics techniques, to verify if the game objectives are appropriate and were accomplished by the users that tested the game.

The structure of the paper is as follows: Section 2 presents related work on serious games for individuals with disabilities. Section 3 briefly describes our game *Downtown, A Subway Adventure*. In section 4 we explain the approach used for the evaluation of the design and development of the game, describing each of the stages. An analysis of the data obtained through the traces collected is shown as a result in section 5. Section 6 summarizes the outcomes from the data analysis and the conclusions obtained.

II. RELATED WORK

Previous research has addressed different applications of learning analytics in serious games. However, few research studies have investigated the needs of individuals with

cognitive disabilities [17]. The work of [18] reviewed the effects of serious games on people with intellectual disabilities or autism spectrum disorder and found that games had a positive impact on the players.

The games tested with users with disabilities were mainly computer serious games; we have not found many studies where more advanced technologies (e.g. augmented reality) are applied to players with intellectual disabilities [19]. From our experience, the use of physical immersive devices in users with intellectual disabilities is discouraged by their trainers and educators, as most of them get anxious and uncomfortable when they are touched (especially these with Down syndrome or Autism) and most of them present limitations in the communication capabilities [12], which leads to a lack of concentration when playing the game.

We didn't find any research that can be comparable with *Downtown* in its educational purpose (training ID users in Subway transportation) or typology of targeted users (wide range of users with ID like Down syndrome, ASD or mild cognitive disability).

Regarding the analysis to be performed on the collected data, the sample size restricts the application of artificial intelligence techniques as they require a high amount of data. However, comparing our work with the applications of data science to game and learning analytics data in general contexts [20], we consider that the sample size reached for participants with intellectual disabilities is highly representative. The number of students with ID in each classroom is substantially lower compared with other ordinary educational environments. With larger samples, artificial intelligence techniques could be adequate to improve the game, but their application will increase the already high costs and delay the gathering timelines.

III. DOWNTOWN, A SUBWAY ADVENTURE

Downtown, A Subway Adventure is a 3D realistic graphic



Figure 1. The game is designed realistically to promote the learning transfer to reality. Snapshot of the interior of a wagon in *Downtown* (left) and the real wagon in the subway of Madrid (right)

adventure game specifically designed for players with ID, like Down syndrome, certain types of ASD (Autistic Spectrum Disorder) or Mild Cognitive disabilities. The aim of the game is to train the students in moving around the city using the public subway system to promote their autonomy, improving their independent life.

The game was designed as a tool for trainers in the transportation program instructed in Down Madrid, one of the biggest educational associations for ID adults and their families in Spain. Madrid subway network is simulated in the game in a 3D realistic perspective (Figure 1), to help users in transferring the in-game experience to the real world when they are traveling by their own. When playing *Downtown* the users must travel around the subway map finding objects and reaching specific destinations. While the user is traveling, different tasks and events automatically pop-up. These tasks not only train the users but also promote the acquisition of other skills that educators find useful to promote independency, like eye-hand coordination, short-term memory, language skills or spatial memory. The game is also enriched with other situations that are more difficult to train in real life, as how to respond to unwanted social interactions with other travelers or how to deal with access machines malfunctions.

Next section explains in detail the objectives, the tasks proposed and the mechanics of the game as well as the process followed to design and develop *Downtown*.

IV. DESIGN AND DEVELOPMENT PROCESS OF DOWNTOWN

To address the several issues of designing and developing serious games for users with intellectual disabilities, we propose an evidence-based approach driven by Game Analytics data. We have tested this approach in the design and development process of the serious game *Downtown*. The process is divided in four stages as depicted in Figure 2.

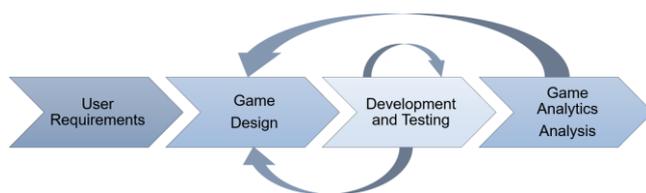


Figure 2. Design and development process cycle used in *Downtown*, A Subway Adventure

To assure that both the educational purpose and the playability of the game are represented in the design and then adequately transferred into the game implementation, we involved in our design and development process three main stakeholders: (1) psychologists and trainers, which main role is to assure the adequacy of the game and its mechanics to the disabled users' cognitive features and abilities (2) researchers and developers, which role is to assure the

playability and accessibility of the learning game and (3) ID users, which role is to early and continuously test the game and provide feedback about the learning experience. We consider that including these actors in the design and development process is a must that, even adding some complexity to the development, in the medium/long run can help to optimize and reduce the total cost of a serious games. Next, we describe each of the stages of the process and their incomes/outcomes.

A. OBTAINING USER REQUIREMENTS FROM EXPERTS

For designing *Downtown*, we interviewed four experts (psychologists and trainers) from Down Madrid to identify the main barriers affecting the capacity of the users about learning concepts with a videogame and other general issues that may affect to their ability for traveling in the subway independently.

The trainers provided a list of aspects that ID students typically find difficult to, like sustaining attention in one particular activity during long periods of time (their attention becomes easily dispersed so any distraction or event that is not directly related with the task which they are working on can scatter it), understanding sequential instructions given at the same time, addressing new problems (even though they could be similar to others than they solved before), executing tasks under time pressure or persisting if don't receive constant feedback [21] [22] [23]. These cognitive characteristics, among others, shaped and delimited the design of *Downtown* since the very beginning of the process.

The game design was guided by these characteristics and included features to fulfill these requirements, for instance, dynamics were included to sustain attention, instructions were constantly kept in screen for users to be able to go back to them easily (Figure 3), no time constraints were set for any part of the game, and constant feedback was provided to players.

Trainers also asked researchers for specific in-game tasks, finally implemented as minigames, to help them improve problem-solving strategies and other skills that can be useful when traveling independently. In particular, minigames should train spatial vision, short-term memory, numerical sequencing and eye-hand coordination.

B. GAME DESIGN

Game Design can be described as the process of designing the rules, the story, the content and the mechanics of a videogame [24]. Usually, this is a creative phase where designers have freedom to create a compelling and engaging game universe. But when developing a game for ID users some of the options are limited by the restrains of their capabilities, as previously explained.

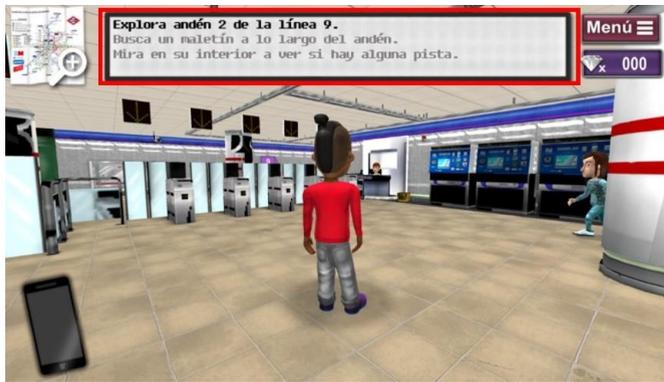


Figure 3. The game asks for tasks sequentially and remain in the screen to help the user remind what the ongoing duty is

First thing in the game design stage is to define the main objectives of the game and the mechanics associated to them. In particular, *Downtown* was designed to accomplish two main educational objectives:

1. Complete a random subway route assigned by the game, changing trains when required
2. Improve the abilities identified by the trainers, required to travel independently: spatial vision, short-term memory, numerical skills and eye-hand coordination)

According to Hunicke [25] a mechanic describes the particular components and dynamics of the game, at the level of data representation and algorithms. Mechanics define the game user interaction and should take into account the specific ID user barriers identified in the requirements phase.

In *Downtown* the user must help the police following and capturing a robber that is traveling in the subway. The game assigns a random starting and ending station and the user must complete the route changing trains when required. There are four levels (easy, medium, hard and expert) that can be selected by the user or the trainer. The length of the route and the number of transfers are related to the difficulty level, from no transfers in easy level to three transfers in the expert level.

Downtown map is designed as a sandbox scenario where users can explore and travel freely through the Madrid Metro map but each game session is programmed to last between 20 and 40 minutes, based on the level of difficulty. This means that the users do not have time pressure while completing the tasks, but educators and researchers can use this observable to analyze the performance of the users while playing.

During the progress of a game session, the game automatically assigns tasks accordingly to the level of difficulty selected. The higher the level, the more difficult the tasks are. Contextual instructions will be given to the user to

find a correct route. Each instruction proposed by the game launch a different problem that the user needs to solve in order to complete the game.

There are also secondary in-game tasks that pop-up during the game that are designed as minigames inside the main plot of the game. These tasks were designed following the specific requirements of the trainers, as described in the previous subsection. Their purpose is not only to sustain the engagement of the users offering new challenges, but also to train practical skills that may be useful when traveling in the subway by their own. These minigames make the progress of the game more dynamic and fun, maintaining the whole game flow.

The resolution time of the minigames and other observables are also tracked and analyzed to evaluate the performance of the users while resolving specific tasks, but users do not have any visual hint of the time spent as we also want to avoid the time pressure element in those minigames.

The main mechanics of the four minigames are described in Table 1.

TABLE I
MAIN MECHANICS ASSOCIATED TO THE MINIGAMES THAT POP-UP DURING A GAME SESSION

Minigame	Skill	Mechanic
Message in a puzzle	Spatial vision	The user has to rotate and sort the pieces of a puzzle to complete a "coded" message. The message gives the user the name of the next station to travel to. The number of pieces in the puzzle increases with the difficulty level.
Simon	Short-term memory	A button-like machine appears in the screen and gives the user a sequence of lights. The user has to remember a sequence of colors and the order of the lights to complete the minigame. The length of the sequence depends on the difficulty level (from 3 colors in the easy level to 6 colors in the expert level)
The safe box	Numerical sequencing	The user finds a safe box in a metro station. To open it and get the clue about the next destination the user has to remember a sequence of numbers. The length of the sequence depends on the difficulty level (from 3 numbers in the easy level to 6 numbers in the expert level)
Shoot your camera	Eye-hand coordination	Once the user reaches the final destination, the game asks for a picture of the robbers. The number of people that should appear in the picture depends on the difficulty level (from 1 character in the easy level to 4 characters in the expert level)

C. DEVELOPMENT AND TESTING

Downtown is a complex game due to its features, learning goals and target players that we have already discussed. We applied SCRUM as the development methodology to have flexibility for testing the game mechanics at the end of each development cycle. This method allowed us to analyze if the initial design was playable to the users' motor and cognitive skills. Down Madrid ID users' beta-tested *Downtown* game six times before releasing the final version, playing parts of the game as they were developed (formative evaluation observed by experts). This feedback provided useful information about the playability and the user experience, reporting what initial game design decisions worked or didn't work. After releasing the final version of *Downtown*, the game was re-engineered to collect the relevant user interaction data using game analytics, so researchers can evaluate the effectiveness of the game.

1) METHODOLOGY

The final version of the game was formally evaluated in an actual classroom. A quasi-experimental design was used to evaluate if the game objectives were achievable by the users and therefore, the game design was adequate to their skills and abilities. The testing sessions took place in the facilities of Down Madrid (Spain). The testing sessions were performed with the totality of the Down Madrid students that attend to the technology class: Fifty-one (51) adults, ages between 19 and 41, with diverse types of intellectual disabilities played the game for three one-hour sessions of effective playing ($n=51$, $M_{age}=29$, $SD_{age}=7.07$). *Downtown* has four difficulty levels. All the players started playing in the easy level and were allowed to start the next difficulty level once they completed all the routes and minigames proposed by the game.

Users were randomly divided in six groups, depending on the schedule availability of each student. Individuals in each group had different IQ, cognitive competences and autonomy. Thirty users were Down (58.8%) while Twenty-one (41.2%) had another type of ID like, Mild Cognitive disability or certain types of ASD.



Figure 4. Users playing *Downtown* during a testing session

All the players played *Downtown* a total of 3 hours. Both trainers and researchers conducted the training sessions, giving advice and helping the users when needed (Figure 4). Parents, guardians and the users themselves were informed about the study and approved the data gathering by signing a consent form.

All the data captured was anonymized to guarantee the privacy of the users. Only trainers, not researchers, were able to match the data with the students for later assessment purposes.

Initial experiment results about gameplays, engagement and the general motivation of the students playing with the game and a more detailed explanation about the methodology used were published in [26], where interaction data was analyzed to refute or validate some of the assumptions that educators expected from the game sessions (e.g., user performance depending on their degree of functionality or user attachment with the main character).

D. GAME ANALYTICS ANALYSIS

One of the major problems that trainers find while working with ID users is the difficulty obtaining reliable feedback about the learning experience of the students. Communication problems are frequent in people with ID, so traditional game evaluation methods, like tests or questionnaires, are not as reliable as when used with neurotypical students [27].

Using evidence-based methods can help both educators and researchers to address this issue. Educators can benefit by obtaining near real-time game user interaction data about the performance of the students while they are playing a game (e.g., they can monitor when users get stuck in one task, check their number of fails vs wins, etc.). Researchers need to validate the decisions made in the game design understanding the capabilities of the users [28], the restrictions that may apply to the design and the learning goals that the trainers expect users to achieve with the game [29].

To address this issue, we applied Game Analytics techniques that can be defined as the process of analyzing the interaction data in serious games, trying to obtain relevant information about the users' behavior in the game (and their learning process in the case of learning games in what has been called game learning analytics [30]).

Downtown game includes a tracker module for sending out relevant in-game information about the users' gameplay interactions. The tracker is open source and has been developed as part of the H2020 RAGE project [31]. All the information captured follows the standard xAPI (Experience

Application Programming Interface), a new specification for collecting, storing and reporting user interactions on learning systems [13]. All this analytics data gathered during the gameplay was sent to a cloud-based data analytics server that provides analyses and dashboards visualizations in near-real time. Using the data analytics services, trainers and researchers were monitoring the users' performance during the game sessions, allowing to identify students that were struggling and helping them when required.

1) IN-GAME METRICS

The tracker captured timestamped information about all the user interactions within the game. Every time an event occurs (e.g., starting/ending a game session, reaching a station, starting/ending a minigame or interacting with the interface elements), the tracker sends out the information to an online server that provides dashboards and metrics visualizations. With all the interaction information, the following metrics were used to perform an analysis to verify if the objectives of the game were accomplished, and therefore validate the design:

- **Total game session time** referred to the average total time spent in one level, from the selection of the character to the completion of the final mission. We consider that one level is complete once a user reaches the final station proposed by the game. If users are able to complete an entire game session, means that they are understanding the mechanics and that they are following the correct route to reach the destination.
- **Average time completing routes** is the time that the users spend traveling in the metro wagon. This time also includes the time spent changing trains when a transfer is required.
- **Inactivity time:** When the user is not progressing in the game (e.g. minigames are not launching because the user is not completing tasks or when the user is clicking randomly in the interface buttons) we consider that the user is inactive. Evidence of high inactivity times in a particular task or level may suggest that the users are having troubles understanding or resolving a mechanic. The inactivity time can also give researchers insights about the engagement of the users in the tasks proposed by the game. Long periods of inactivity time may suggest a passive attitude of the user towards the game and so, a deficit of attention.
- **Overall minigame performance** referred to the number of attempts to resolve a minigame and the total time spent in completing it.

- **Other observables:** like the use of the accessibility options, the number of clicks in the help button or how many times the user consults the metro map.

Using all the in-game data analytics gathered, we created specific constructs to validate the design and the development process of *Downtown* but we consider that this approach can be generalized to other game designs for ID users. Through these constructs we can analyze the data obtained and identify design anomalies. If no anomalies are found in the data analysis, we consider that the game design is adequate for the users. If we find inconsistencies in the data analyzed, next step would be to diagnose the design problem.

V. RESULTS

This section provides an overview about the data analysis conducted at the end of the design and development process. The purpose of the analysis is to verify if the main objectives of *Downtown*, and therefore its game design, is adequate and achievable by the users. More than 163.000 data points regarding the users' interaction data were included in the analysis. Both trainers and researchers took part in the process: data collected were analyzed by the researchers and it serves to validate or refute the impressions of the trainers about the interaction of the users with the game. The purpose of this process is to confirm if the game can be used as a learning tool. It is also an example of how to use Game Analytics to validate a game design and its mechanics, identifying errors.

A. TOTAL GAME SESSION TIME

At the beginning of the game design phase, experts believed that not all the users were going to be able to complete the routes proposed by the game, which is the main objective of *Downtown*. To verify this hypothesis, we analyzed the average total game time and the average number of incorrect stations transited in a game session.

92.15% of the users ($n=47$) were able to complete at least one correct route during the game sessions. Only four Down users were not able to reach any destination ($n=4$, 7.85%). More than 72.8% of the incorrect paths occurred in the first 30 minutes of playing which suggest that the users needed to play the game at least once to fully understand the mechanics of the game and be familiar with the interface (note that the users played the game for three hours). This result was not consistent with the initial experts' expectations in the game design phase but confirms that both the tutorial of the game and the difficulty levels were adequate for the users.

Trainers also expected that the ID users that play videogames in a regular basis as part of their daily life entertainment ($n=21$, 41.1%) would complete the tasks faster and would do less mistakes than the rest of the users. Only 10.56% of the

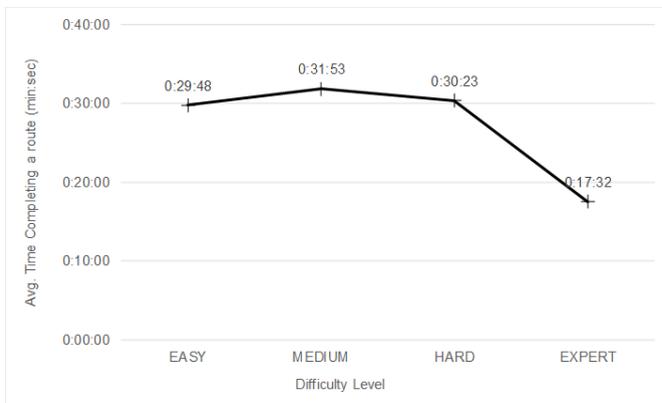


Figure 5. Average total session time per difficulty level
 routes followed by those ID players were wrong, compared to a 20.37% of the non-players routes ($\chi^2 = 4.3761$ p -value = 0.0364 < 0.05. The results are statistically significant). The average time completing tasks for ID players was 28:18 minutes, compared to the non-players that spent an average of 32:16 minutes per game session which confirms the initial opinion of the trainers.

B. AVERAGE TIME COMPLETING ROUTES

The experts also aimed to verify if the complexity of the difficulty levels was well-adjusted to the users' cognitive skills, so they could adapt and customize the game sessions to the capabilities of each user in the classroom. We analyzed the average time completing routes of all users per difficulty level to validate the design of the game levels.

In average, users spent 30:41 minutes completing a route in the levels easy, medium and hard. The expert level is not significant for the analysis as there were only five users that reached the expert level ($n=5$). All these users were regular players and their capacities were above the average, according to the trainers.

When analyzing this observable, we found that it takes more time to complete the medium level than the easy and hard level (Figure 5). As stated before, *Downtown* automatically assigns random starting and ending stations, depending on the level of difficulty that the users are playing. In the easy level, the route is direct, which means that there are no transfers during the journey. The number of transfers increases progressively with the levels: one transfer in medium level, two in hard and three in the expert level. Thus, the higher the level, the more time that the user needs to reach the final destination.

Researchers find two possible explanations for this abnormality: (1) the tasks are not appropriate for the users' cognitive abilities, so they are getting stuck in the medium level trying to complete it or (2) there is a design/development flaw in the game when assigning the routes.

C. AVERAGE INACTIVITY TIME

One method to verify if the design of the tasks is adequate to the users is to analyze the average inactivity time (Figure 6). If the users are playing the game and are not 'inactive' means that they are understanding the mechanics proposed by the game.

This metric decreases as the users reach higher levels, from an average of 00:55 seconds in the easiest level to 00:32 seconds (-41.8%) in hard and expert levels (Figure 6) and is not co-related with the average time completing routes. This data is in accordance with the initial expectations of the trainers: even though the complexity of the routes proposed by the game increases as the difficulty level is higher, the users are more 'active' because they have previously learned the mechanics of the game and understand better how to resolve the puzzles. Trainers also noticed that the users asked less for help as the difficulty level of the game increases.

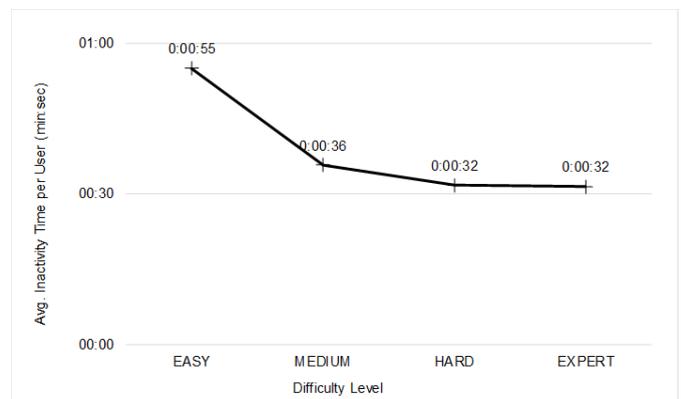


Figure 6. Average inactivity time per difficulty level

The analysis suggests that the design of the tasks is in accordance with the users' competences.

D. AVERAGE NUMBER OF STATIONS TRANSITED

As data showed no abnormal behavior of the average inactivity time, we analyzed the average number of stations that the users transited during a game session, sorted by level (Table 4) and found that the number of stations is higher in medium level than in hard and expert level. Also, this number remains the same in hard and expert level.

This evidence indicates that the game is not accurately selecting the optimal routes for the levels to be balanced what can be considered a flaw in the development of the game.

Researchers think that even though almost all the users completed the routes that the game asked, and so the objective of the game can be considered as achieved from an observational perspective, the game analysis uncovered a design flaw that is making the users complete the paths slower than they should. This bug could have been difficult

to detect using formative evaluation or other observational-based methods.

TABLE II
AVERAGE NUMBER OF STATIONS TRANSITED IN A GAME SESSION SORTED BY DIFFICULTY LEVEL

Difficulty level	Average number of stations transited per route
Easy	5
Medium	11
Hard	10
Expert	10

E. OVERALL MINIGAME PERFORMANCE

Trainers and researchers designed four minigames to train four skills that can be useful when traveling in the subway. These minigames are embedded in the main plot so they break the routine of the game and promote the user engagement (Figure 7). By analyzing the average completion time of each minigame, experts could evaluate the competences of the users related to each skill.

From a general perspective, users spent more time and did more mistakes in the Simon minigame (short-term memory) and in the in-game camera minigame (eye-hand coordination) and obtained better results in the puzzle (spatial vision) and safe box (numerical sequencing) minigames (Figure 7).



Figure 7. Screenshots of the minigames (puzzle, Simon, safe box, camera)

In the puzzle minigame all the users put the pieces in their place without mistakes, even though the number of pieces increased with the difficulty level. The average resolution time remained constant in all levels (01:32 minutes) which is considered normal by the experts, as the users played puzzles with the computer before.

The resolution time of the safe box minigame remains constant between levels too (01:23 minutes). Only two users out of Twenty-seven (7,4%) failed in remembering the

sequence of numbers in the safe box minigame. Both were Down and non-regular players.

Almost all users (77.7%) did more than one try before completing the camera minigame ($\bar{X}=2.02$). While two tries are not considered a high number by the trainers, the data analysis shows a standard deviation of the completion time of almost one minute (00:57), which means that for some users it took almost twice the time to resolve the minigame compared to the rest of the classroom. Data didn't show a pattern about the type of ID or the frequency of playing videogames of the players that spent more time in this minigame.

The same scenario happened in the Simon minigame where users had to remember a sequence of color lights in the right order. Almost all of the users (91.3%) tried almost three different sequences before inserting the right combination ($\bar{X}=2.57$, $\sigma=2.93$), regardless the level that they were playing in. Users spent an average of 2:07 minutes, the highest resolution time of all minigames. The data analysis shows a standard deviation of the completion time of 01:15 minutes. In this case, almost the 70% of the users that spent less time than the average in completing the minigame were regular videogame players, which suggests that playing videogames can help in enhancing the short-term memory.

According to the data, the difficulty of the Simon and the camera minigame is higher than expected by the trainers, which means that either (1) the design of the tasks is not completely adequate to the users' cognitive skills or (2) the abilities that these two minigames promote (eye-hand coordination and short-term memory) are not as developed as the other two in this sample. In both cases these two minigames unexpectedly increased the total complexity of *Downtown*.

Nevertheless, the total number of mistakes done by the users in this two last minigames decreases as the users reached higher levels. This circumstance shows an improvement of the users' performance at higher levels, so it seems that users are able to complete the tasks related to eye-hand coordination and short-term memory after practicing, despite their type of disability or previous experience with videogames.

VI. CONCLUSION

In this paper we aimed to exemplify an approach that uses Game Analytics data to evaluate the design and development of a learning game for users with ID. Applying game analytics techniques researchers can investigate if the in-game observables provide useful information about the effectiveness of the game design and if the objectives of the game are adequate to the users' cognitive abilities. This approach has been applied to *Downtown*, *A Subway*

Adventure, a game that aims to train users with ID in traveling in the subway independently. However, we consider that the process described is generalizable to validate other learning games and, in particular, in domains where the opinion of the users is difficult to obtain or not even fully reliable (as it is the case with ID users).

First step of the process is the user requirements stage where we interviewed the experts from Down Madrid to understand the cognitive restraints related with the use of videogames that ID students have. Next, in the game design stage, researchers established the main objectives of the game and the mechanics associated. In particular, *Downtown* goal is to complete a random route assigned by the game and train the abilities required to travel in the subway independently. We tested the game in the facilities of Down Madrid with 51 users with ID. Trainers and researchers conducted the training sessions capturing in-game data about the interaction of the users with the game.

After capturing the data, researchers and trainers performed an analysis to verify if the users were able to achieve the objectives of the game, validating the game design.

Almost all of the users were able to complete the routes that the game proposed, making less mistakes after the first thirty minutes of gameplay. Also, users that play videogames in a regular basis did less mistakes and completed the tasks faster than non-players.

A flaw in the game implementation was hidden behind these promising results: the analysis of game analytics data showed an error in the game mechanics. *Downtown* was assigning longer routes in the medium level than in the advance one. This mistake was hidden to other traditional approaches (e.g. beta-testing, observational-based methods).

Regarding the results related to the minigames focused in training the abilities to travel independently, from a general perspective, the data analysis shows a good performance in spatial vision and numerical sequencing. However, users struggled in the minigames that train eye-hand coordination (camera) and short-term memory (Simon) despite their type of disability. From the data gathered, we cannot determine if these results uncover a problem in the game design of the minigames or if our users find harder to improve these skills. To get detailed information about this issue it would be necessary to re-capture and analyze new data, which is now unfeasible as the development is completed. This situation could be avoided if the game analytics approach were used also with the mini-games in the early game testing.

Next step in our research will be conducting a case study with a group of users that played the videogame and evaluate their competences when traveling in the subway in real life. We will later compare their performance with users that were

not previously trained with the game to verify the effectiveness of the game design and test the transfer of knowledge from the game to reality.

ACKNOWLEDGMENT

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. We also want to thank Rodrigo Ribeiro-Pinto Carvajal for his excellent work developing *Downtown*. 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 P2018/TCS4307), by the Ministry of Education (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

- [1] D. R. Michael and S. L. Chen (2005). Serious Games: Games That Educate, Train, and Inform. Muska & Lipman/Premier-Trade.
- [2] T. M. Connolly, E. a. Boyle, E. MacArthur, T. Hainey, and J. M. Boyle (2012). A systematic literature review of empirical evidence on computer games and serious games. *Computers and Education*, vol. 59, no. 2, pp. 661–686, Sep.
- [3] Boyle, E. A., Hainey, T., Connolly, T. M., Gray, G., Earp, J., Ott, M., Pereira, J. (2016). An update to the systematic literature review of empirical evidence of the impacts and outcomes of computer games and serious games. *Computers and Education*, 94, 178–192. <https://doi.org/10.1016/j.compedu.2015.11.003>
- [4] P. Wouters, C. van Nimwegen, H. van Oostendorp, and E. D. van der Spek (2013). A meta-analysis of the cognitive and motivational effects of serious games. *Journal of Educational Psychology*, vol. 105, pp. 249–265, 2013.
- [5] S. Freeman et al., (2014). Active learning increases student performance in science, engineering, and mathematics, *Proceedings of the National Academy of Sciences*, vol. 111, no. 23, pp. 8410–8415.
- [6] R. Prieto de Lope, N. Medina-Medina, P. Paderewski, and F. L. Gutiérrez-Vela. (2015). Design methodology for educational games based on interactive screenplays, in *Cosecivi, Congreso de la Sociedad Española para las Ciencias del Videjuego*.
- [7] G. Petri, and C. Gresse von Wangenheim (2016). How to Evaluate Educational Games: a Systematic Literature Review. *Journal of Universal Computer Science*, 22(7), 992–1021.
- [8] A. Rigaki, A. Ntagianta, A. Leonidis, D. Ioannidi, M. Korozi, and C. Stephanidis. (2019). Learning Analytics for Ami Educational Games Targeting Children With Cognitive Disabilities. *INTED2019 Proc.*, vol. 1, no. March, pp. 8916–8926, 2019.
- [9] American Psychiatric Association, *Diagnostic and statistical manual of mental disorders (5th ed.)*. (2013). Arlington, VA: American Psychiatric Publishing.
- [10] C. González-Ferreras, D. Escudero-Mancebo, M. Corrales-Astorgano, L. Aguilar-Cuevas, and V. Flores-Lucas. (2017). Engaging Adolescents with Down Syndrome in an Educational Video Game. *International Journal of Human-Computer Interaction*, vol. 33, no. 9, pp. 693–712.
- [11] M. Terras, E. Boyle, J. Ramsay and D. Jarrett. (2018). The opportunities and challenges of serious games for people with an intellectual disability. *British Journal of Educational Technology*. 49. 690-700. 10.1111/bjet.12638.

- [12] S. Lanfranchi, C. Cornoldi, and R. Vianello. (2004). Verbal and Visuospatial Working Memory Deficits in Children With Down Syndrome. *Am. Journal of Mental Retardation*, vol. 109, no. 6, pp. 456–466, 2004.
- [13] S. Tsikinas, S. Xinogalos, M. Satratzemi, and L. Kartasidou (2011) Using Serious Games for Promoting Blended Learning for People with Intellectual Disabilities and Autism: Literature vs Reality.
- [14] V. Hulusic and N. Pistoljevic. (2017). A curriculum for developing serious games for children with autism: A success story. 2017 9th International Conference on Virtual Worlds and Games for Serious Applications (VS-Games) - Proc., no. September, pp. 149–152.
- [15] P. Kato and S. de Klerk. (2017). Serious Games for Assessment: Welcome to the Jungle,” *Journal of Applied Testing Technology*, vol. 18, no. S1, pp. 1–6.
- [16] Á. Serrano-Laguna, I. Martínez-Ortiz, J. Haag, D. Regan, A. Johnson, and B. Fernández-Manjón. (2017). Applying standards to systematize learning analytics in serious games. *Computer Standard Interfaces*, vol. 50, no. June 2016, pp. 116–123.
- [17] A. Nguyen, L. A. Gardner and D. Sheridan (2018). A framework for applying learning analytics in serious games for people with intellectual disabilities. *British Journal of Educational Technology*, 49(4), 673–689. <https://doi.org/10.1111/bjet.12625>
- [18] S. Tsikinas and S. Xinogalos. (2018). Studying the effects of computer serious games on people with intellectual disabilities or autism spectrum disorder: A systematic literature review. *Journal of Computer Assisted Learning*. 10.1111/jcal.12311.
- [19] A. Rigaki, A. Ntagianta, A. Leonidis, D. Ioannidi, M. Korozi and C. Stephanidis (2019). Learning Analytics For AMI Educational Games Targeting Children With Cognitive Disabilities. *Proceedings of INTED2019 Conference 11th-13th March 2019, Valencia, Spain*. 8916-8926. 10.21125/inted.2019.2220.
- [20] C. Alonso-Fernández, A. Calvo-Morata, M. Freire, I. Martinez-Ortiz, B. Fernández-Manjón. (2019). Applications of data science to game learning analytics data: a systematic literature review. *Computers and Education*, vol. 141. <https://doi.org/10.1016/j.compedu.2019.103612>
- [21] A. Djuric-Zdravkovic, M. Japundza-Milislavjevic, D. Macesic-Petrovic, and A. Djuric-Zdravkovic. (2010). Attention in children with intellectual disabilities. *Procedia - Social and Behavioral Sciences*, vol. 5, no. 2, pp. 1601–1606.
- [22] H. E. Kirk, K. Gray, D. M. Riby, and K. M. Cornish. (2015). Cognitive training as a resolution for early executive function difficulties in children with intellectual disabilities. *Research in Developmental Disabilities*, vol. 38, pp. 145–160.
- [23] M. M. Rahman, S. M. Ferdous, S. I. Ahmed, and A. Anwar. (2011). Speech development of autistic children by interactive computer games. *Interactive Technology and Smart Education*, vol. 8, no. 4, pp. 208–223.
- [24] S. Rogers, *Level Up! The guide to great video game design*, 2nd ed. Wiley (2014).
- [25] R. Hunnicke, M. LeBlanc, and R. Zubek. (2004). MDA: A Formal Approach to Game Design and Game Research. *AAAI Workshop - Technical Report*, pp. 1–4.
- [26] A. R. Cano, B. Fernández-Manjón, and Á. J. García-Tejedor. (2018). Using game learning analytics for validating the design of a learning game for adults with intellectual disabilities. *British Journal of Educational Technology*, vol. 49, no. 4, pp. 659–672.
- [27] J. L. McLeskey, M. S. Rosenberg, and D. L. Westling, *Inclusion: Effective Practices for All Students*, 2 edition. Pearson (2012).
- [28] M. W. Rodrigues, L. E. Zárate, and S. Isotani. (2018). Educational Data Mining: A review of evaluation process in the e-learning. *Telematics and Informatics*.
- [29] D. Hicks, M. Eagle, E. Rowe, J. Asbell-Clarke, T. Edwards and T. Barnes (2016). Using game analytics to evaluate puzzle design and level progression in a serious game. *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge - LAK '16*, 440–448. <https://doi.org/10.1145/2883851.2883953>
- [30] M. Freire, Á. Serrano-Laguna, B. Manero, I. Martínez-Ortiz, P. Moreno-Ger, and B. Fernández-Manjón. (2016). *Game Learning Analytics: Learning Analytics for Serious Games*. Learning, Design and Technology (pp. 1-29). Cham: Springer International Publishing. http://doi.org/10.1007/978-3-319-17727-4_21-1.
- [31] “RAGE Realizing an Applied Gaming Eco-System.” [Online]. Available: <http://rageproject.eu/>.



ANA RUS CANO was born in Madrid, Spain, in 1983 and lives in Chicago, IL, since 2014. She received her B.S. degree in computer science from the University Francisco de Vitoria, Madrid, in 2009. She also received a M.S. degree in new technologies and animation, from the University Francisco de Vitoria, Madrid, in 2009, a M.S. degree in business advisory, from EOI (industrial organization school), Madrid, in 2011 and a M.S. degree in marketing of the entertainment business, from ESIC business school, Madrid, in 2011. She is a PhD candidate since 2014 in the e-UCM e-learning research group at UCM, Madrid. She has worked for advisory firms like IBM, PwC and Prisa as a senior consultant and business developer, in technology, media and entertainment departments.

Her research interest is focused on game design and learning games, and more specific, in serious game developments for users with intellectual disabilities. Her latest articles are focused on the application of learning analytics, game learning analytics and data analysis to improve the design of serious games.



ÁLVARO J. GARCÍA-TEJEDOR Ph.D. in Biochemistry and passionate about computer science, he has been linked to the University since 1.994. He has been Lecturer in UCM, UC3M and Universidad Francisco de Vitoria, where he is now Associate Professor in Artificial Intelligence. He has made compatible for fifteen years his university profile with his work in Consulting and Software Development companies (Indra and Telefonica Data, among others), involved in more than 20 R&D projects mainly in Artificial Intelligence (11 of them in European Programs), leading National and European teams and directing development and innovation departments. He has also been appointed as external consultant by the European Commission for several Research Framework Programs (FP5 and 6). As a researcher, he started doing mathematical modeling of biological systems. Now his areas of interest are bio-inspired computing (neurocomputing, evolutionary computation), philosophy of AI and serious games. His current work is focused on neural models, deep learning techniques and their applications in several fields. He especially contributes to the practice of eHealth and social care by exploiting massive biomedical data for remote care, diagnostic and optimization.



CRISTINA ALONSO FERNÁNDEZ obtained her Bachelor in Computer Science and her Bachelor in Mathematics for the Complutense University of Madrid in 2016. A year later, she finished the Master in Data Mining and Business Intelligence, also from the UCM. Since September 2016, she is part of e-UCM. She has worked as a researcher of the H2020 Beaconing Project until October 2018. Her research interests include educational videogames and application of data analytics and data mining for their improvement.

Learning Analytics techniques and standards are one of her main research areas.



BALTASAR FERNANDEZ-MANJON received the PhD degree in physics from the Universidad Complutense de Madrid in 1996. He is a Full Professor of computer science at UCM and director of the e-UCM e-learning research group. He has the Honorary Complutense-Telefonica Chair on Digital Education and Serious Games. He is the co-author of two books, more than 200 articles His research interest is focused on the applications of ICT in education and in serious

games and educational simulations applied to different domains (e.g. medicine, education). He is also working in game learning analytics and the application of e-learning standards to the integration of games in e-learning systems. He is a senior member of the IEEE.