This document is published in:

Manuel J. Gomez, José A. Ruipérez-Valiente, Pedro A. Martinez, and Yoon Jeon Kim. 2020. Exploring the Affordances of Sequence Mining in Educational Games. In Eighth International Conference on Technological Ecosystems for Enhancing Multiculturality (TEEM'20). Association for Computing Machinery, New York, NY, USA, 648–654.

DOI: https://doi.org/10.1145/3434780.3436562

Exploring the Affordances of Sequence Mining in Educational Games

Manuel J. Gomez
Universidad de Murcia
Murcia, Spain
manueljesus.gomezm@um.es

José A. Ruipérez-Valiente
Universidad de Murcia and MIT
Murcia, Spain
jruiperez@um.es

Pedro A. Martinez
Universidad de Murcia
Murcia, Spain
pedroantonio.martinezs@um.es

Yoon Jeon Kim
MIT
Cambridge (MA), USA
yjk7@mit.edu

ABSTRACT

Games have become one of the most popular mediums across cultures and ages and the use of educational games is growing. There is ample evidence that supports the benefits of using games for learning and assessment. However, we do not usually find games incorporated into educational environments. One of the main problems that teachers face is to actually know how students are interacting with the game as they cannot analyze properly the effect of the activity on the students. To improve this issue, we can use the data generated by the interaction of students with such educational games to analyze the sequences and errors by transforming raw data into meaningful sequences that are interpretable and actionable for teachers. In this study we use a data collection from our game *Shadowspect* and implement learning analytics with process and sequence mining techniques to generate two metrics that aim to help teachers make proper assessment and better understand the process.

CCS CONCEPTS

•Applied computing → Education; E-learning; • Software and its engineering → Interactive games.

KEYWORDS

Educational games, learning analytics, game-based assessment, sequence mining

ACM Reference format:

Manuel J. Gomez, José A. Ruipérez-Valiente, Pedro A. Martinez and Yoon Jeon Kim. 2020. Exploring the Affordances of Sequence Mining in Educational Games. In *Proceedings of the Eight International Conference on Technological Ecosystem for Enhancing Multiculturality (TEEM'20). ACM, New York, NY, USA, 7 pages.* https://doi.org/10.1145/1234567890

1 INTRODUCTION

Digital games have become significant part of families and young people around the world. According to a recent survey conducted in U.S. [7], early three-quarters (74%) of American parents believe that video games can be educational for their children, and more than half (57%) enjoy playing games with their children at least weekly. In the past 10 years, numerous studies (see [3] for a meta-analysis) have reported that games can be more effective than other traditional teaching methods to support both content learning and skill development, and many studies report teachers' positive perceptions regarding use of games in classrooms to promote students' engagement, motivating, and learning [14]. Despite the benefits of educational games as well as teachers' positive perceptions, actual implementation of game-based curriculum in schools remains still limited. A 2014 survey of teachers [8] reports that 57% of the responded teachers use games at least once a week, but they are largely, 33% of the respondents, unsure about how integrate game with the curriculum. That is why providing guidelines and facilitating a simplified deployment of these games is so important, so that their implementation can greatly benefit teachers and students [1].

One instructional challenge that teachers face when they implement the game in their curriculum is that they can't gain a quick sense of who is progressing productively, who is confused and disengaged, or who is persisting unproductively. That is, unlike non-digital curricular materials such as worksheets, teachers can't easily identify different students' needs by just hovering over their gameplay. Additionally, using data generated from gameplay requires data literacy skills and understanding the reliability and origin of the data, which many teachers might not have sufficient training and familiarity about [20]. One way to support a better integration of digital games with regular curriculum is to provide interactive and intuitive data visualization tools that can help teachers to make quick instructional decisions using a series of interactions of the student with the environment.

Due the highly interactive and open-ended nature of the game environment, it presents rich opportunities to assess how students are learning beyond how well they are doing [16, 23, 26]. For example, Spring and Pellegrino [12] discuss how the open-ended nature of the game environment can lead to development conceptual knowledge as well as learning through failure. However, these open-ended environments also present some challenges. Open-ended game structure allows players to perform a wide variety of actions. The actions of the players can easily become misaligned with respect the original objectives, because there are more chances for the game to present conflicting feedback when players have more freedom of action [11]. Another difficulty we usually find in games is to know and understand players' actions within the game. Therefore, making meaningful use of the complex and numerous data generated from gameplay can benefit from the practices of the learning analytics field [26].

Learning analytics is the measurement, collection, analysis and reporting of data about learners and their contexts, for purposes of understanding and optimizing learning and the environments in which it occurs [21]. Analyzing sequences allow us to examine the specific processes students perform while interacting with a learning environment, and not just outcome performance measures [33]. A potential opportunity to give teachers the information they need, might lie in analyzing students' sequences and errors, and making them available for teachers, so that they can make proper assessment and better understand the learning process.

In this paper, we report two sequence mining metrics that could be actionable and interpretable for teachers: sequences of actions within a puzzle and common errors related to the puzzle solution. We use *Shadowspect*, which is a 3D geometry computer game where students can create geometric primitives such as cones or spheres to solve 3D puzzles, developing their geometric, dimensional and spatial reasoning skills. Using *Shadowspect*, large amounts of raw data are collected to analyze and transform them into useful information so that teachers can better understand students' interaction with the game. We also present a set of visualization prototypes for teachers that communicate the metrics in an intuitive and actionable way. More specifically, this paper addresses the following objectives:

(1) To present a proposal of sequence mining metrics: one to analyze the sequences of actions performed by students and another one to analyze their most common errors by puzzle.
(2) To present a case study with uses cases from data collected in K12 schools across the US using *Shadowspect.* This case study includes visualizations for teachers that exemplify how to interpret these metrics and visualization to better understand students' behavior with the game and intervene.

The rest of the paper is organized as follows. Section 2 reviews background literature on the applications of process and sequence mining, learning analytics and development of visualizations in educational environments. Section 3 describes the methods, including *Shadowspect* as well as the data collection. Section 4 presents the definition of the two metrics. Next, Section 5 describes different use cases using the metrics and visualizations developed. Then, we finalize the paper with discussion in Section 6 and conclusions and future work in Section 7.

2 RELATED WORK

The use of virtual open-ended environments (e.g. games or simulations) in education present new opportunities to support student learning. The information obtained can be used not only to help teachers to manage their classes, understand their students' learning processes, and reflect on their own teaching methods, but also to support learners' self-awareness of their own actions and to provide feedback to learners [28]. Whatever goal they might have, for simulations to be successful they need structural elements to give them shape, and this often comes from the rules of game-play and/or digital enhancements [27]. Students need guided facilitation from the teacher in order to be effective instructionally in the classroom. However, the beauty of simulations is that they create learning opportunities and experiences that might otherwise never be able to be created in the traditional classroom [19].

A game (educational or not) can generate vast amounts of interaction data, even in a short game-play session. The application of data mining and visualization techniques on player interaction logs can provide very valuable insights to game developers regarding how players are interacting with the game[9]. However, the difficulties on measuring learning outcomes achieved through serious games have been a main barrier for successful deployment and adoption of serious games within formal education [13].

A possibility to analyze these data generated is to use process and sequence mining. They can be used in different environments, but in this research, we use them applied in education. Process mining allows us to examine the specific processes students follow during learning [17]. There are many different types of sequence mining techniques, such as sequential pattern mining, differential sequence mining or process mining [33]. Although sequence and process mining are two related areas, their objectives are not exactly the same. The objective in process mining is to discover underlying processes from data, and the objective in sequence mining is to find common patterns between data examples where the values are delivered in a sequence [2]. As an example, in [18] authors study transformation of action sequences using action features, such as activity categorizations, relevance and timing between actions, and repetition of analogous actions. As raw data can contain a large amount of information which is not related to the students' interaction, it is very important to transform this data into an appropriate sequence of actions. In this study, we will analyze students' sequences of actions while playing the different puzzles included in the game.

Once information is obtained from analyzing the data, the next step is to make it easy to understand and present it in an easily actionable way. This will help people involved in the learning process (e.g. teachers, students or other stakeholders) to understand easily the information provided. Graphic visualizations of these data can be used to support this type of analysis [22]. Visualizations are one of the most important components of research presentation and communication, because of their ability to represent large amounts of data [34] and also because it is easier for the brain to comprehend an image versus words or numbers [4]. In this research, we use visualizations to graphically represent the information we have obtained from our sequence mining metrics. When creating visualizations, it is important to follow the guideline proposed in [15]: create the simplest graph that conveys the information you want to convey. A large variety of visualizations can be used in data mining [10]: 2D and 3D scatter plots, heat maps, polar charts, contour plots... Some other researchers have used visualizations to represent information obtained from data [6, 25, 30, 31] across different educational contexts. For example, in the area of games for learning, authors in [24, 29] also have developed visualizations to simply represent data from students' interaction with games. In this work we go a step further, presenting a new approach using sequence mining in an educational game. Since it is difficult to find sequence works focused on games, we present a strong opportunity of how to use sequence mining in a game to infer useful information and make it actionable via visualizations.

Figure 1: Two puzzles examples in *Shadowspect*.

3 METHODS

In this section, we introduce the educational game, context and data collection of this research.

3.1 Shadowspect

We use *Shadowspect*, a game-based assessment tool that aims to provide metrics related to geometry content and other behavioral and cognitive constructs. *Shadowspect* has been designed explicitly as a formative assessment tool to measure math content standards (e.g. visualize relationships between 2D and 3D objects), thus teachers can use it in their core math curriculum.

When students begin a puzzle, they receive a set of silhouettes from different views that represent the figure they need to create, which will be composed of other primitive shapes the student can put into the scenario. The primitive shapes that students can create are cubes, pyramids, ramps, cylinders, cones and spheres. Depending on the level and difficulty, the puzzle may restrict the quantity or type of shapes they can create. After putting these shapes in the scenario, they can also scale, move and rotate the shapes in order to build a figure that solves the puzzle. Students can move the camera to see the figure they are building from different perspectives and then use the 'Snapshot' functionality to generate the silhouette and see how close they are to the objective. Finally, they can submit the puzzle and the game will evaluate the solution and provide them with feedback[1].

In the version of *Shadowspect* that we have used in this work, we have 9 tutorial, 9 intermediate and 12 advanced levels. The tutorial levels aim to teach the basic functionality of the game, so the students can learn how to build different primitives, scale and rotate them, how to change the perspective, take snapshots and so on. The intermediate levels allow students more freedom so they will not receive so much help to solve puzzles and then the advanced levels pretend to be a real challenge for students who have gained experience with previous levels before. This set of levels provides a linear sequence of increasing difficulty puzzles. However, students can move from any puzzle to another, regardless of its difficulty or order.

3.2 Educational Context and Data Collection

The data used for this paper was collected as part of the initial data collection to build the assessment machinery of *Shadowspect*. The team recruited seven teachers who used the game in their classes for at least two hours in theory math classes (class grade from 7th grade to 10th). In this paper, we represent a real case scenario of how the teacher of a class could use these visualizations to monitor the progress of their students

[1] See a demo online: https://youtu.be/j1w_bOvFNzM

in the classroom. All student interactions with the game were collected and stored in a MySQL database, and we do not collect any identifiable or personal data from the users except for a nickname provided by themselves. The complete data collection from a total of 322 students includes around 428,000 events (an average of 1,320 events per user). Students were active in the game environment for 260 hours (an average of 0.82 active hours per student), and students solved a total of 3,802 puzzles (an average of 13 puzzles per student).

4 SEQUENCE MINING METRICS PROPOSAL

In this Section, we present our two sequence mining metric proposals: one of them related with the sequence of actions within each puzzle attempt and the other is related to common errors in the solving process. The first metric aims to obtain the actions performed by students within a puzzle while playing in order to analyze them and observe possible problems and solutions to them. The second metric provides a way to automatically identify the most common errors for each puzzle based on the information obtained in the previous metric.

4.1 Sequences Within Puzzles

In this metric, the objective is to obtain a sequence of actions for every puzzle attempted by each student, so that we can reconstruct the low-level actions performed while playing *Shadowspect*. This metric is divided in the following two main steps:

(1) **Data Transformation**: We transform the raw data into an adequate sequence of actions that are representable. This step also includes data cleaning to keep only useful events, in this case we only keep those events related to the puzzle solving process: starting a puzzle, manipulation events (create, delete, scale, rotate or move a shape), snapshots, perspective changes and puzzles checks.
(2) **Data Compacting:** We reduce the number of events but maintaining the information that is needed for building a sequence of actions. We compact those events that are the same by adding an additional field that indicates the number of times that an event has been performed in a row. For example, if the student has changed the perspective of the game three times in a row, the original data containing three different events will be transformed in a single perspective change event that has been performed three times. When the event is related to the manipulation of shapes, we only compact them if they are related to the same shape identifier.

We reduce significantly the amount of data: from 428,000 events in the original data collection to 107,000 (i.e. 25% less events) after processing the metric. With this metric we obtain a detailed sequence of the actions each student conducted while trying to solve a puzzle, and at the same time we reduce the amount of data making it easier to understand.

4.2 Common Errors

This metric uses the output obtained from the sequences within puzzles and provides a way to identify common errors in the resolution of *Shadowspect* puzzles. Therefore, this allows teachers to quickly locate the errors committed by the students and focus on clarifying those aspects to improve the learning process.

One initial detail to explain is that each puzzle might have multiple solutions. The game auto-solver retrieves the silhouettes of the "master" solution, and checks if the silhouettes of the current figure submitted by the student match the master solution. Hence, there might be multiple figures that have the same silhouettes from the master solution, however, this is not extremely common: In ten of the puzzles, shapes provided by the student were equal to the ones in the master solution in 80% of cases. Another group of seven puzzles present identical shapes in 50% of cases, and then there is another set of two puzzles which has identical shapes in only 5% of the cases. As the solutions can be different, it is more difficult to find what is an actual error instead of a student trying to solve the puzzle in an alternative way. To solve this issue, we have to make an assumption: we will analyze only puzzles solved using the same shapes of the master solution. In first place, we apply the sequences within puzzles metric to obtain the initial input that this metric uses. Then, we apply an algorithm that has the following two steps:

(1) **Identify meaningful events**: We identify the changes a student has made in the shapes between a failed submission and a correct submission. For example, if a student submits a puzzle and the solution is incorrect, and then the student creates a pyramid and deletes a cone in the scenario, those edits are registered by our algorithm as changes between submits.
(2) **Compute most common errors**: Once we have registered all those changes after wrong submissions, we group them by puzzle to obtain the shapes and manipulation events that the students have had problems within each puzzle.

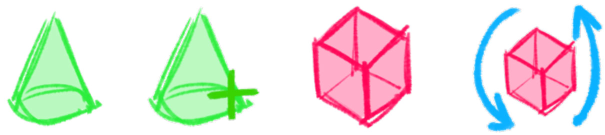
This metric can greatly facilitate error finding for teachers implementing *Shadowspec*t in a class.

5 CASE STUDY

This section presents the case study that exemplifies how teachers can visualize these metrics to assess different situations. In the case study, we first describe the visualization design, then we have a use case for sequences within puzzles and another one for common errors.

5.1 Visualization Design

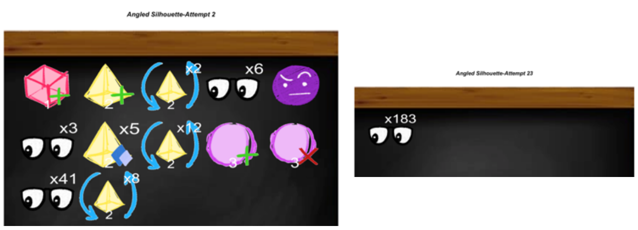
We have developed some icons that match the different shapes and manipulation events that a student can make through the game, in order to generate representative visualizations.



(a) (b) (c) (d)

Figure 2: Examples of icons developed.

In Figure 2 we can see some examples of the primitive images (Figure 2a and 2c) that have been used to create composite images taking into account the possible actions that can be made with each shape. For example, in Figure 2b we can see how the addition of a cone is represented. Another example is presented in Figure 2d, with the icon exemplifying the rotation of a cube.



(a) (b)

Figure 3: Two examples of sequences within puzzles.

In Figure 3 we can see two different examples of the sequences within puzzles visualization for two different student attempts on the same puzzle. First of all, let's explain how the visualization works. For every single action a student can make, we have created an image that represents graphically that action. For example, in Figure 3a we can see that the first three actions are the creation of a cube, a pyramid and then the rotation of a pyramid. To represent the number of times an action has been performed, we indicate the number of times in the upper-right corner. To represent each type of shape a student can interact with, we use different graphics for each type of shape (cones, ramps…), and then we indicate the shape identifier in the puzzle with a number in the bottom-center.

In both sub-figures we see interesting behaviors. In Figure 3a, in the final sequence of actions we see the student creates a sphere, delete it and then changes the perspective 41 times in a row. Finally, the student rotates a pyramid three times and then leaves the puzzle, abandoning the intent to solve the puzzle. On the other side, in Figure 3b we see that the only action the student has made has been changing the perspective 183 times, showing a clearly off-task behavior. That way, the teacher can see in detail every action the student has made in the solving process of a puzzle and really assess if they were trying to complete the puzzle or if they are clicking buttons randomly (like in Figure 3b).

5.2 Sequences within Puzzles in "Square Cross-Sections"

Now we are going to see how two different students solve the same puzzle with different sequences of actions.

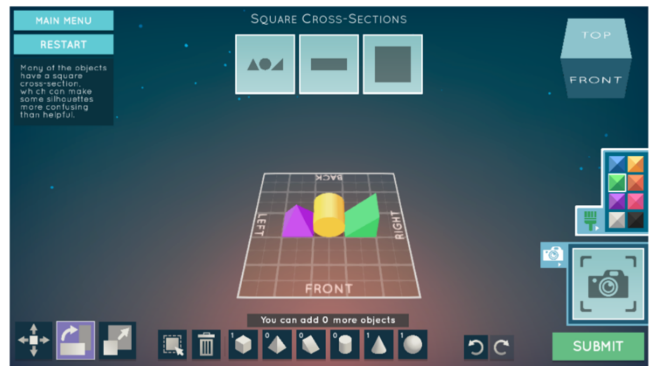


Figure 4: Square Cross-Sections example in *Shadowspect.*

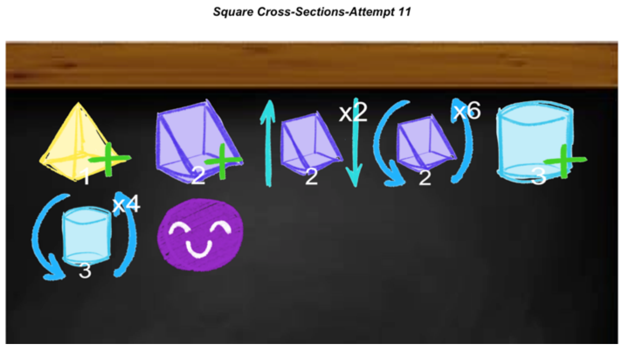
In Figure 4 we see an example of how we can solve "Square Cross-Sections" puzzle in *Shadowspect.* As we can observe, the puzzle can be solved using a pyramid, a cylinder and a ramp. Let's now see a student's sequence solving this puzzle.

In Figure 5a we can see that the student has solved the puzzle in a few number of events. After creating a pyramid and a ramp, the student moves and rotates the ramp to match the views required and then he creates a cylinder and rotates it. That way, the student solves the puzzle without making any incorrect submits, showing confidence on its actions.

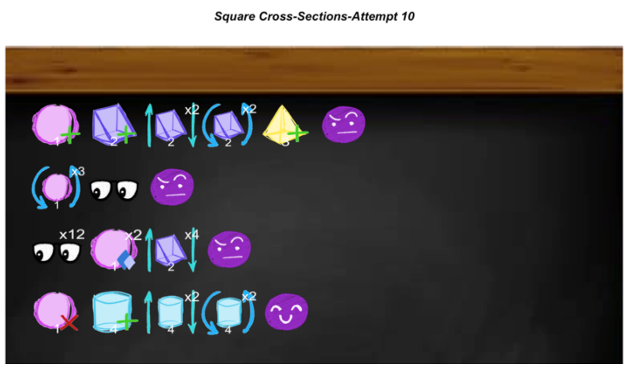
In Figure 5b we now see a student solving the same puzzle but with a different sequence of actions. The first thing we notice is that the student creates a sphere instead of a cylinder. This will do the job with the first silhouette showed in Figure 4, but it does not match with the other two silhouettes. After doing an incorrect submit, the student tries to change the rotation of the sphere and its size, but the solution is still incorrect. Then the student realizes the sphere is not the adequate shape and he deletes it and creates a cylinder instead to solve the puzzle correctly.

5.3 Common errors in "Square Cross-Sections"

After seeing this sequence of actions related to the puzzle, a teacher may want to know the common errors that his class of students are having in this specific puzzle. To do that, the teacher can use the common errors metric and visualization and locate easily the most common incorrect actions made by students.



(a)



(b)

Figure 5: Two examples of sequences within puzzles in "Square Cross-Sections".

In Figure 6, we see an example of the common errors visualization. It has two different parts: On the top we can see the shapes that compose the master solution and, on the bottom, we see the most common errors and its percentage. In this specific group, we see that the most common errors are related to the creation and rotation of cylinders and the deletion of spheres. This matches perfectly with the sequence we have analyzed for the previous student, where the student used a sphere instead of a cylinder and then he had to delete it and create a new cylinder to solve the puzzle.

6 DISCUSSION

At the beginning of this research we defined the context and main objectives of our work. The two metrics defined provide new useful information that can help teachers in their classes, providing critical insights to make an in-depth monitoring of students while playing *Shadowspect*. Then, with the visualizations developed, teachers can easily understand metrics output, evaluate their students and easily intervene to solve possible problems when necessary. We had to make an assumption when implementing the common errors metric, since puzzles do not have a unique solution. Because of that, we have made a limited implementation only taking into account puzzles that have been solved with the master solution. We do not have a validation with teachers yet, thought this is part of the future work as we are co-designing a dashboard with a cohort of teachers.

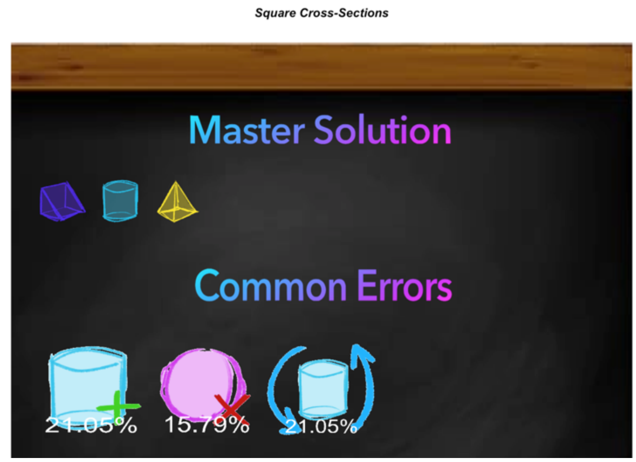


Figure 6: "Square Cross-Sections" common errors.

In this work we try to explore the potential of using temporal sequences to analyze student's interaction, instead of only using static numbers after playing the game. That way, we are analyzing student's behavior over time. Other researchers have used sequence and process mining previously to analyze student's behavior in games. In [32] sequence mining is used to assess how metacognitive monitoring and scientific reasoning impacted the efficiency of game completion during learning using a game-based learning environment. We can see other examples that also have used process mining, as in [5], where authors use process and sequence mining with predictive purposes. Our approach is very teacher-centered, clearly targeting the application of sequences within a classroom, but based on this previous research, we could also consider using process mining to predict students' outcomes in puzzles.

Using the potential of these sequences, teachers can use the common errors metric and visualization in their groups to detect misconceptions and revise them in the class or with individual students. This also provides a new approach to personalized support in the classroom class, helping to adapt the pedagogy to the current modern times with the support of technology.

7 CONCLUSIONS

The objective of this research was twofold: First, to propose two metrics that can provide detailed information regarding the process of students with the puzzles in *Shadowspect* and their errors. Second, to achieve simple but detailed visualizations of these two metrics that can allow teachers to monitor students within their classes, so they can evaluate the students' performance or detect common errors quickly and effectively.

This approach can help alleviate one of the main problems when implementing educational games in the classroom, which is the possibility to better understand how students interact with the game and to locate the problems that the students are having while playing. This also represents an opportunity for educators to provide personalized attention to their students and help them in their learning process.

The next step in our work is to implement just-in-time interventions that aim to provide support at the right time, adapting to the need of each student. We are also working on the co-creation of a dashboard for teachers that can provide greater speed and interactivity when displaying the visualizations and data. This will enable just-in-time interventions during sessions. Also, we will be working on obtaining evidences of the interpretability of these visualizations and to make them explainable so that teachers can easily intervene. We are currently developing new metrics, related to play styles in the game and persistence. More nuanced metrics and visualizations will allow students to visualize their mistakes and areas of improvement. In this way we can use *Shadowspect* as a robust learning tool with that can be easily implemented by teachers in the classroom and that emphasizes the formative feedback to the student.

ACKNOWLEDGMENTS

We want to acknowledge support from the MIT-SPAIN "la Caixa" Foundation SEED FUND and the Spanish Ministry of Economy and Competitiveness through the Juan de la Cierva Formación program (FJCI-2017-34926).

REFERENCES

[1] Alonso-Fernández, C. et al. 2018. Improving serious games analyzing learning analytics data: lessons learned. *International Conference on Games and Learning Alliance* (2018), 287–296.

[2] Bogarín, A. et al. 2018. A survey on educational process mining. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery.* 8, 1 (2018), e1230.

[3] Clark, D.B. et al. 2016. Digital games, design, and learning: A systematic review and meta-analysis. *Review of educational research.* 86, 1 (2016), 79–122.

[4] Cukier, K. 2010. A special report on managing information. *The Economist.* 394, 8671 (2010), 3–18.

[5] Deeva, G. et al. 2017. Dropout prediction in MOOCs: a comparison between process and sequence mining. *International Conference on Business Process Management* (2017), 243–255.

[6] Duval, E. 2011. Attention please! Learning analytics for visualization and recommendation. *Proceedings of the 1st international conference on learning analytics and knowledge* (2011), 9–17.

[7] ESA 2019. *2019 Essential Facts About the Computer and Video Game Industry.*

[8] Fishman, B. et al. 2014. Empowering educators: Supporting student progress in the classroom with digital games. *Ann Arbor, MI: University of Michigan. Retrieved from http://gamesandlearning. umich. edu/agames.* (2014).

[9] Freire, M. et al. 2016. Game learning analytics: learning analytics for serious games. *Learning, design, and technology.* Springer Nature Switzerland AG. 1–29.

[10] Grinstein, U.M.F.G.G. and Wierse, A. 2002. *Information visualization in data mining and knowledge discovery.* Morgan Kaufmann.

[11] Harpstead, E. et al. 2015. Replay analysis in open-ended educational games. *Serious games analytics.* Springer. 381–399.

[12] Harpstead, E. et al. 2015. Replay Analysis in Open-Ended Educational Games. 381–399.

[13] Hauge, J.B. et al. 2014. Implications of learning analytics for serious game design. *2014 IEEE 14th international conference on advanced learning technologies* (2014), 230–232.

[14] Huizenga, J.C. et al. 2017. Teacher perceptions of the value of game-based learning in secondary education. *Computers & Education.* 110, (2017), 105–115.

[15] Kelleher, C. and Wagener, T. 2011. Ten guidelines for effective data visualization in scientific publications. *Environmental Modelling & Software.* 26, 6 (2011), 822–827.

[16] Kim, Y.J. and Ifenthaler, D. 2019. Game-based assessment: The past ten years and moving forward. *Game-Based Assessment Revisited.* Springer. 3–11.

[17] Kinnebrew, J.S. et al. 2013. A contextualized, differential sequence mining method to derive students' learning behavior patterns. *Journal of Educational Data Mining.* 5, 1 (2013), 190–219.

[18] Kinnebrew, J.S. and Biswas, G. 2012. Identifying Learning Behaviors by Contextualizing Differential Sequence Mining with Action Features and Performance Evolution. *International Educational Data Mining Society.* (2012).

[19] Klopfer, E. et al. 2009. Using the technology of today in the classroom today: The instructional power of digital games, social networking, simulations and how teachers can leverage them. *The Education Arcade.* 1, (2009), 20.

[20] Kukulska-Hulme, A. et al. 2020. Innovating Pedagogy 2020: Open University Innovation Report 8. (2020).

[21] LAK 2011. LAK 2011 : 1st International Conference Learning Analytics and Knowledge.

[22] Leony, D. et al. 2012. GLASS: a learning analytics visualization tool. *Proceedings of the 2nd international conference on learning analytics and knowledge* (2012), 162–163.

[23] Loh, C.S. et al. 2015. Serious games analytics: Theoretical framework. *Serious games analytics.* Springer. 3–29.

[24] Martínez, P.A. et al. 2020. Visualizing Educational Game Data: A Case Study of Visualizations to Support Teachers. (2020).

[25] Minović, M. and Milovanović, M. 2013. Real-time learning analytics in educational games. *Proceedings of the first international conference on technological ecosystem for enhancing multiculturality* (2013), 245–251.

[26] Owen, V.E. and Baker, R.S. 2020. Learning analytics for games. *Handbook of game-based learning. Cambridge, MA: MIT.* (2020).

[27] Prensky, M. 2003. Digital game-based learning. *Computers in Entertainment (CIE).* 1, 1 (2003), 21.

[28] Romero, C. and Ventura, S. 2010. Educational data mining: a review of the state of the art. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews).* 40, 6 (2010), 601–618.

[29] Ruiperez-Valiente, J.A. et al. 2020. Patterns of engagement in an educational massive multiplayer online game: A multidimensional view. *IEEE Transactions on Learning Technologies.* (2020).

[30] Ruiperez-Valiente, J.A. et al. 2016. Scaling to massiveness with analyse: A learning analytics tool for open edx. *IEEE Transactions on Human-Machine Systems.* 47, 6 (2016), 909–914.

[31] Ruipérez-Valiente, J.A. et al. 2015. ALAS-KA: A learning analytics extension for better understanding the learning process in the Khan Academy platform. *Computers in Human Behavior.* 47, (2015), 139–148.

[32] Taub, M. et al. 2018. Using sequence mining to reveal the efficiency in scientific reasoning during STEM learning with a game-based learning environment. *Learning and instruction.* 54, (2018), 93–103.

[33] Taub, M. and Azevedo, R. 2018. Using Sequence Mining to Analyze Metacognitive Monitoring and Scientific Inquiry Based on Levels of Efficiency and Emotions during Game-Based Learning. *Journal of Educational Data Mining.* 10, 3 (2018), 1–26.

[34] Ware, C. 2019. *Information visualization: perception for design.* Morgan Kaufmann.