Exploring Play-Learners' Analytics in a Serious Game for Literacy Improvement

Shamus P. Smith[1(✉)], Daniel Hickmott[2], Erica Southgate[2],
Ross Bille[1], and Liz Stephens[2]

[1] School of Electrical Engineering and Computer Science, The University of Newcastle,
Callaghan, NSW 2308, Australia
{Shamus.Smith,Ross.J.Bille}@newcastle.edu.au
[2] School of Education, The University of Newcastle, Callaghan, NSW 2308, Australia
{Daniel.Hickmott,Erica.Southgate}@newcastle.edu.au

Abstract. The collection and analysis of analytics incorporated into serious games provides researchers with objective data on player behavior related to serious game design elements and learning. Such analytics offer insights about play-learners engagement that is not possible to capture through traditional techniques. Visualization of learning behavior data can allow for a comparison between the pedagogical intent of the game design and the play-learners actual behavior within the serious game. This paper describes the use of game-play logs to identify pathways through gaming content in a serious game app for literacy improvement. The paper describes the technical aspects of processing game-play logs and their transformation into visualizations, and considers how these visualizations can be used to explore play-learner behavior in relation to the pedagogical intent of activities embedded in the serious game app.

Keywords: Game-play analytics · Serious game · Literacy app · Visualization · Pedagogy · Gamification · Learning

1 Introduction

The collection and analysis of analytics incorporated into serious games provides researchers with objective data on player behavior related to serious game design elements and learning. Such analytics offer insights about play-learners engagement in serious games that is not possible to capture through traditional techniques [11]. There is increasing interest in the use of game-play logs and the transformation of this data through visualization techniques, as a means of exploring evidence on effectiveness [11, 13]. Visualization of learning behavior data can allow for a comparison between the pedagogical intent of the game design and the play-learners actual behavior within the serious game. The purpose of this paper is to: (i) describe the technical aspects of processing game-play logs; (ii) account for the transformation of game-play logs into visualizations that have scalable form; and, (iii) to use these visualizations to explore play-learner behavior in relation to the pedagogical intent of learning activities within a serious game app. These three areas will be examined in the context of an evaluation

conducted on a literacy improvement app called *Apostrophe Power*. For this paper, three questions relating to play-learners' behavior during game-play were explored. These were:

– In which areas of the game do play-learners spend most of their time?
– What order do play-learners complete the game levels?
– After making mistakes, do play-learners choose to view the training videos?

This paper is structured in the following way. Section 2 describes the serious game app, *Apostrophe Power*. Section 3 provides a brief overview on work related to serious game analytics and use of game-play logs. Section 4 outlines preparation of the logs and the pipeline for visualization. Section 5 provides three cases of game-play log visualizations and interprets play-learner behaviors in the context of the pedagogical intent of the game. Section 6 offers suggestions on the use of game-play logs and visualizations, and discusses their implications for learning design in serious games.

2   *Apostrophe Power*: A Literacy App

The *Apostrophe Power* serious game is an app that students can use to improve their apostrophe use. Lukač [12] notes that "apostrophe 'misuse' is the most popular topic in the field of language advice", with misuse common in social media communication attributed to incorrect usage of apostrophes [6, 12]. The app focused on the main functions of apostrophes, referred to as categories:

– *Contractions* - for example, the contraction of cannot is can't.
– *Ownership* - for example, singular and plural ownership, e.g. Dad's car.
– *Misfits* - the appropriate use of "it's" and "its".
– *Combination* - exercises that combine the use of the previous three categories.

These categories were used to organize learning exercises in the game. *Apostrophe Power* provides learners with a series of increasingly difficult unlockable levels across these four categories. In addition, short instructional training videos on these categories were also provided and could be returned to during the game to scaffold learners towards mastery [7]. Events in the game were logged when play-learners interacted with any of the game components. An example of a game level in *Apostrophe Power* can be seen in Fig. 1. In an *Apostrophe Power* level the play-learner's goal is to get a mouse character across sinking platforms, from left to right, to a prize (a block of cheese) on the final platform.

To get to the next platform, the play-learner corrects sentences on the bottom half of the screen and taps the Jump button. If the sentence has been corrected successfully, the mouse will jump to the next platform. Each level has ten platforms, so the play-learner must correct ten sentences to complete a level. The play-learner places apostrophes, by tapping their finger in the appropriate positions to make the sentence correct. The platforms begin to sink once the mouse has landed on them, adding time pressure to the game. If the platform sinks into the water the play-learner has to restart the level. An incorrect answer will result in a popup dialog that offers play-learners the choice of

another question or the option to view the training video associated with the current category, or to return to the level selection screen.

Figure 2 shows the possible navigation paths through the app between the different areas of the game. The main motivation for the analysis of the game-play logs was to discover how play-learners navigated through and interacted with the learning activities.
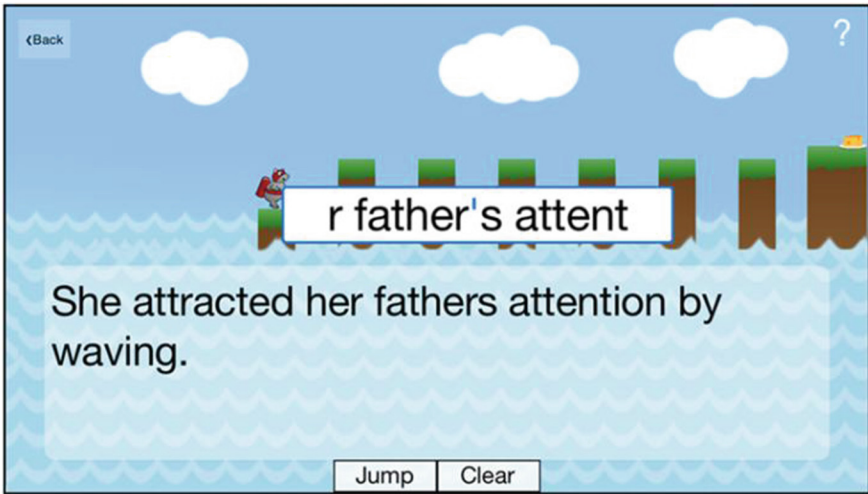

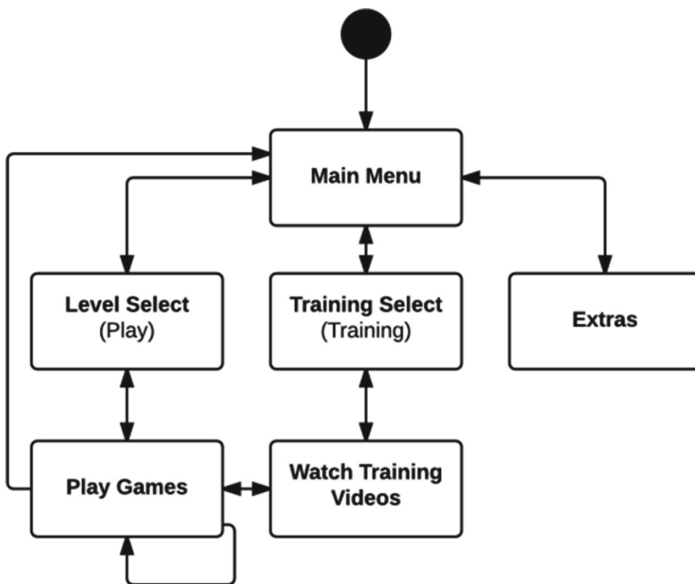**Fig. 1.** Game screen of *Apostrophe Power*

Fig. 2. Possible navigation paths within the *Apostrophe Power* app

3   Related Work

According to Loh, Sheng and Ifenthalder [11, p. 22], the primary purposes of serious games analytics are to "obtain valuable actionable insights to better the game or learning design", and to "improve skills and performance of the play-learners to better convince stakeholders of the game's effectiveness". The collection methods for data used for serious games analytics can be separated into two categories: *in-situ* and *ex-situ* [11]. In-situ collection occurs in the game itself (for example, logging game-play events), whereas ex-situ is data collected outside of the game. Focus group discussions and post-test surveys are examples of ex-situ data. Barab, Gresalfi and Ingram-Goble [1] assert that the analysis of game-play logs (collected in-situ) can reveal insights about the understanding play-learners have of the concepts covered in the game.

We were unable to find any studies that involved in-situ data collection that were specifically related to serious games that aim to improve adults literacy skills, the target demographic for the *Apostrophe Power* app. Browne, Anand and Gosse [2] designed and evaluated a series of m-learning literacy apps for adults using ex-situ techniques, with pre-test and post-test questionnaires. However, they did not perform any in-situ collection. Snow, Allen and McNamara [15] applied a variety of techniques, including random walks, entropy analysis and Hurst exponents, on game-play logs collected from a game-based system, titled iSTART-2 which involved the teaching of self-explanation and comprehension strategies at a high-school level. These techniques were used to gain an understanding of play-learners' decision and behavior patterns.

In studies that involved serious games analytics with *in-situ* data collection there were some common methods used for discovering patterns in play-learners' decision making behavior. Cluster analysis is often used to profile play-learners, using feature vectors of different metrics derived from game-play logs [3, 8, 10]. The aim is to identify particular patterns or strategies used by play-learners. Bayesian networks are another method of analysis that has been used [4, 10]. Bayesian networks model probabilities of decisions as a graph. In a serious game, the nodes in the resulting graph could be states in the game and the probabilities of certain transitions between these states could be used to gain insight into play-learners' decision making during game-play. Loh and Sheng [10] note that Bayesian network models may be suitable for the educational testing industry, but that the serious games industry should use other, more innovative methods of analysis.

The most widely used technique for analysis of *in-situ* data is visualization. Visualization is often used at the end of serious games analysis [13]. This is because visualizations are an effective way of sharing results as graphical representations can show a large amount of detail in a small amount of space [4]. Dicerbo et al. [4] notes that visualizations can expose relationships more effectively than other methods. Data visualization can also be used to discover patterns in play-learners' use of serious games [9, 16].

The technical details of the processing of the game-play logs and the creation of visualizations are not typically described in the literature. The current work provides a case study of this and focuses on custom software, which used 3rd party libraries where appropriate, to perform the data processing and visualization.

4   The Preparation of Game-Play Logs for Visualization

The process used for analysis of game-play logs can be separated into three main steps: *Pre-Processing*, *Data Transformation* and *Visualization*. The first two steps are explained in this section while the visualization of the results is detailed in Sect. 5. An overview of the process can be seen in Fig. 3, that shows how different parts of the process fit into these three steps.
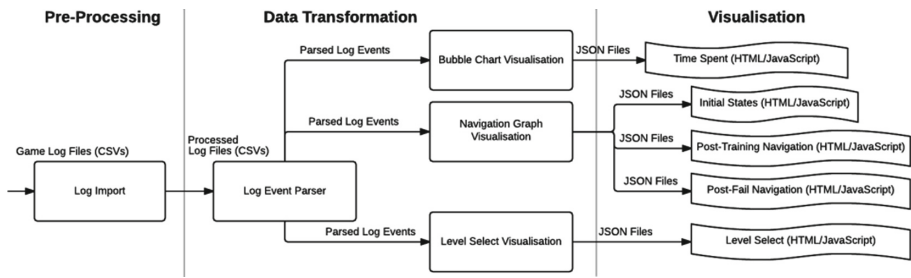

Fig. 3. Game-play log processing pipeline

4.1   Data Collection and Pre-processing

During 2015, a series of *Apostrophe Power* evaluation sessions were conducted at The University of Newcastle, Australia. The participants were provided with Samsung Galaxy Tab3 Android tablets, with the app pre-installed, and were asked to play the game for approximately 45 min. While they played, their interactions with the app were recorded as events to a local SQLite database on the tablet. At the end of the session, the recorded interactions were exported as a CSV (Comma Separated Values) file via email from the tablet.

A total of 17 participants took part in the evaluation sessions. The highest number of events that were logged was 3780 and the lowest was 1692. An example of the contents of the CSV file of game-play events can be seen in Fig. 4. The first step was to write code (the *Log Import* module, shown in Fig. 3) to pre-process these CSV files, changing the structure of the data and removing events that were not needed for the current analysis. The first six lines of the file were removed, as these were either administrative entries, e.g. data separator identifier, the unique device ID, or summary metrics from the device, e.g. the Boolean values of level completion. Next, all of the events that did not take place on the evaluation day were removed. Finally, a row of headers were added, the separators changed to commas and the modified rows exported to a new file.

The exported CSV files were then input into the *Log Event Parser* module (see Fig. 3). This module was responsible for transforming the recorded events into a list of objects representing the different states of the game in the order, and with the duration, that the play-learner had navigated them in. The different states map to different areas of the game, for example *Playing Game* or *Watching Training Video*. The resulting data included the current category and/or level, when relevant to the state. The type of

transition that the play-learner made to the state was also recorded in the event. For example, the play-learner may have navigated to a game level after completing a training video.

The results from this parsing were used in the visualizations described in Sect. 5. An example of the parsed events, from the data in Fig. 4, in CSV file format is shown in Fig. 5.

```
sep=;
30044c60c720b100
ownership;[true, true, true, true]
contractions;[true, true, true, true]
misfits;[true, true, true, true]
mixed;[true, true, true, true]
2015-08-28 07:48:35.448; close_dialog_ClearAppState; ;
2015-08-28 07:48:38.220; close_activity_Settings; ;
2015-09-01 14:02:05.485; open_activity_MainMenu; ;
2015-09-01 14:02:09.147; click_button_Play; ;
2015-09-01 14:02:09.564; open_activity_LevelPicker; ;
2015-09-01 14:02:11.659; click_button_ownershipLevel_0; ownership; 0
2015-09-01 14:02:12.017; open_activity_Game; ownership; 0
2015-09-01 14:02:12.234; load_question(The airports terminal is busy.); ownership; 0
```

Fig. 4. Fragment of a data file from a game-play session

```
Begin, End, StateName, TransitionFrom, Category, Level, DurationSeconds
1/09/2015 2:02:05 PM,1/09/2015 2:02:09 PM,Main Menu,Start,,,3
1/09/2015 2:02:09 PM,1/09/2015 2:02:11 PM,Level Select,From Main Menu,,,2
1/09/2015 2:02:11 PM,1/09/2015 2:03:16 PM,Playing Game,From Game Level Select,ownership,0,65
1/09/2015 2:03:16 PM,1/09/2015 2:04:49 PM,Playing Game,From completed level,ownership,1,92
1/09/2015 2:03:16 PM,1/09/2015 2:06:42 PM,Watching Training Video,From incorrect question,ownership,,205
```

Fig. 5. New data file exported to CSV after performing event parsing

4.2   Choice of Technologies

The code for the analysis of the *Apostrophe Power* game logs had two main purposes: the processing of the logs and the visualization of the transformed data. In this section we explain our reasons of choosing F# for data processing and HTML and JavaScript for visualization.

F# is a functional-first programming language, developed and maintained by the F# Software Foundation and Microsoft (http://fsharp.org/). F# is referred to as a functional-first language. This is because it is possible to use imperative programming constructs, such as while loops, in F#. However, developers are encouraged to use functional constructs as much as possible. Although F# is not a commonly used programming language[1], it provides features that made it an appealing choice for data processing and transformation. Firstly, F# allows compilers to infer types from a sample file in CSV, JSON and other formats without having to write a parser. Thus the transformation of the CSV files was straightforward with no need for new code for parsing the files.

Secondly, F# has higher-order functions and piping syntax for manipulating collections. This makes it possible to approach the design of the data processing and

---

[1] F# is ranked #36 in the TIOBE Index. See http://www.tiobe.com/tiobe_index [last access 10/05/2016], an index of rankings of the popularity of different programming languages.

transformation in a way that Fowler[2] refers to as a *Collection Pipeline*. This aided the process (as seen in Fig. 3) of data transformation and visualization that involved taking collections, performing some manipulation and then sending it to the next step in the processing pipeline. Finally, F# focuses on immutability. Unless specified all variables declared in F# will be immutable. If the data processing and transformation was to be executed in parallel, for example through multi-threading, the use of immutability would make the implementation of this safer [5]. Loh, Sheng and Infenthaler highlight the need for serious games analytic methods to be scalable [11] thus algorithms that can be run in parallel may be an appropriate way for scaling methods.

In this project HTML and JavaScript were used to produce data visualizations for three reasons. Firstly, there are several high-quality JavaScript libraries available for creating visualizations in SVG (Scalable Vector Graphics) format, such as Data-Driven Documents (see *d3.js*, http://www.d3js.org/) and *vis.js* (http://visjs.org/). Secondly, if future work is performed on the project developers will not need to install additional software packages to create visualizations. Developers would only need access to a modern web browser. Thirdly, the researcher working on data analytics of this project was more familiar with these technologies than other data visualization software packages. This accelerated the development of the processing pipeline and the resulting visualizations.

The graphs presented in Sect. 5 were originally produced using the graph visualization software Graphviz (http://www.graphviz.org/). However, using Graphviz would mean that anyone doing work on the project in the future would have to install and configure this software package to produce visualizations. The creation of these graphs was then implemented in HTML and JavaScript instead (by using vis.js). This meant that if a developer wanted to reproduce these visualizations with new data, they will only need to run the F# code and would be able to view the resulting visualizations in any modern web browser. This provides a level of robustness and future proofing to the analytics pipeline described here.
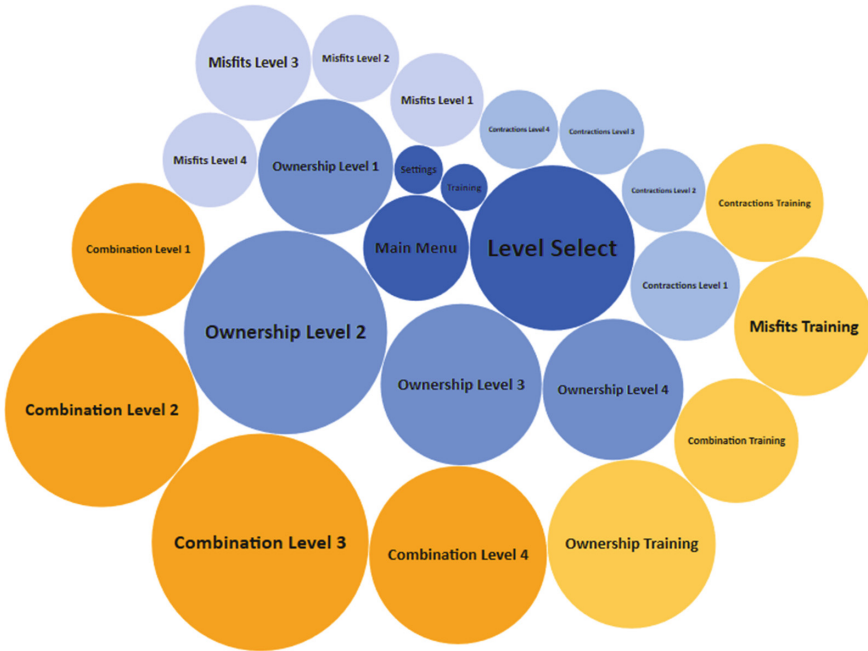
5   Game-Play Log Visualization

In this section we discuss the visualizations generated through the process described in Sect. 4. The aim is to consider questions relating to play-learners' decisions during the *Apostrophe Power* evaluation (see Sect. 1).

5.1   In Which Areas of the Game Do Play-Learners Spend Most of Their Time?

Discovering how much time play-learners spent in different parts of the game is an important metric. If, for example, play-learners spent most of their time outside the training area of the game, this could indicate that the training area was unattractive or that the game was too easy, with implications for learning design.

---

2   See http://martinfowler.com/articles/collection-pipeline/ [last access 11/5/2016].

Fig. 6. Bubble chart of time spent in different areas of the *Apostrophe Power* app

A bubble chart visualization[3] was produced from the parsed game-play logs (see Fig. 6). The bubbles in Fig. 6 with category name, "Level" and a number represent time spent in that level. The bubbles with the category name and suffixed with "Training" represent time spent watching videos. As can be seen in Fig. 6, the participants spent most of their time playing through the game levels. The play-learners spent significantly more time in the Combination category than in any of the other category. This suggests that the play-learners found the questions in this category more difficult than in the other categories. This is consistent with the pedagogical intent of the design where play-learners need to bring together their knowledge of all the functions of apostrophe use to successfully master the Combination levels.

The bubble chart produces a high level view of activity to indicate broad trends. However, one advantage of this visualization is that if more logs were collected, the bubble chart could be produced for the new data with minimal effort. Thus it provides scalability.

5.2   What Order Do Play-Learners Complete the Game Levels?

A visualization (Fig. 7) was created to represent the order that players completed the game levels in. We were interested in both the navigation between levels and how this

---

[3] Source d3.js bubble chart example from https://bl.ocks.org/mbostock/4063269 [last access 11/5/2016].

corresponded with our interface design. Thus the visualization layout has been mapped to the level selection screen (see Fig. 8). Each of the arrows in Fig. 7 represents a transition from one of level to another. The thickness of an arrow indicates how common the transition was. From this visualization, it can be seen that the most common path for completing the game was to navigate the game by completing all the levels in a category and then continuing to the next category.
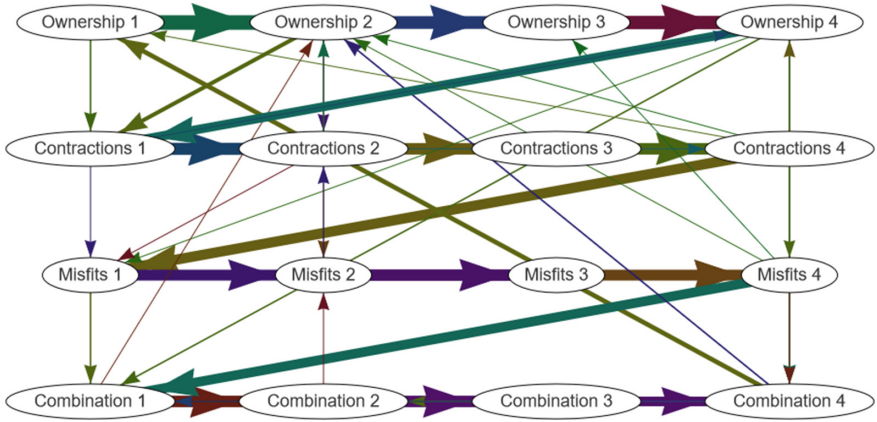
Fig. 7. Visualization of the order that players completed the game levels. The thickness of an arrow indicates how common the transition was with participants
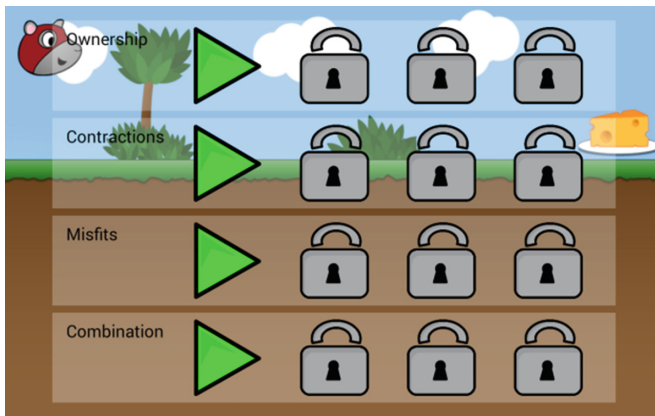
Fig. 8. Level selection screen in *Apostrophe Power*

In addition to completing all the levels in a category from easy to most difficult (left to right) the play-learners also attempted to master single categories of apostrophe use before attempting the levels where apostrophe functions are combined in activities (hence the movement from top to bottom). This was consistent with the pedagogical intent of the design where play-learners needed to bring together their knowledge and

mastery of all the functions of apostrophe use before attempting to complete the Combination levels.

5.3 After Making Mistakes, Do Play-Learners Choose to View the Training Videos?

The transitions after a play-learner made an unsuccessful attempt were classified into three groups: (i) *Level Menu* - the play-learner returns to the Level Select Menu; (ii) *Watch Training Video* - the play-learner chooses "Go To Training", which navigates to the relevant training video; and, (iii) *Try Again* - the play-learner chooses to make another attempt, which will generate a new exercise in the current category. These transitions were combined for all play-learners, and the percentages of these transitions were calculated. This was visualized by creating a state transition diagram, using viz.js (see Fig. 9).

The percentage (1.66 %) of transitions to a training video after making an unsuccessful attempt was very low compared to the percentage (96.46 %) of transitions where play-learners tried again after making an unsuccessful attempt. This suggests that, while there is a pop-up dialog that offers the play-learner the choice to navigate to the training section after an unsuccessful attempt, it could have been more effective to force play-learners to watch videos after a number of unsuccessful attempts.

This is an example where the pedagogical intent in the design did not manifest in the play-learner behavior. Player-learners generally did not use the instructional training videos as a scaffold. Instead, player-learners persisted with repeated unsuccessful attempts. A key question that this raises is how the unsuccessful attempts relate to the different apostrophe categories and levels of difficulties. This requires further investigation.
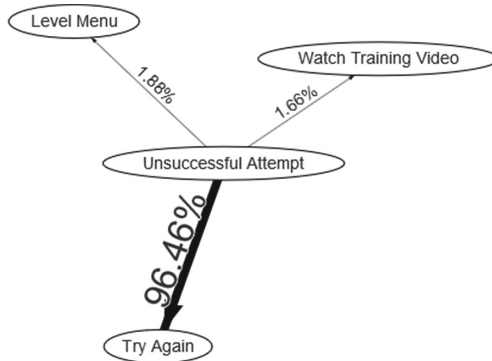

Fig. 9. The pathways play-learners took after unsuccessful attempts

6   Conclusions

This paper has described the use of game-play logs and their visualizations to identify pathways through gaming content in the *Apostrophe Power* app, to elucidate learning

behavior. As the aim of the project was to produce a serious game, mapping the engagement between game-based learning exercises and the associated training videos was useful. It provided evidence on the appropriateness of each component, e.g. the difficulty of the exercises or the value of the training videos. This paper focuses on the value of visualization to conduct such analysis.

This paper outlines the process of taking event logs and processing them into a form to support visualizations for analysis. Specifically, this paper describes the use of F#, HTML and JavaScript to build a scalable pipeline for visualizations. The aim has been to define an environment that is both scalable, with potential support for parallelism, and with future proofed visualization output. In this case the visualizations are accessible via standard web browsers. Thus the methods described here are scalable and maintainable, so that later evaluations could be analyzed using the same methods and code with minimal effort.

Three questions relating to play-learner behavior during game-play were explored. In two cases, namely "In which areas of the game do play-learners spend most of their time" and "What order do play-learners complete the game levels", the visualization confirmed that the play-learners behavior was consistent with the pedagogical intent of the serious game design. However, in the third case - "After making mistakes, do play-learners choose to view the training videos?" - where it was assumed that unsuccessful attempts would prompt player-learners to view an instructional video, there was limited evidence of this choice. Thus, the pedagogical intent did not manifest in the play-learner behavior. This said, findings from the evaluation, reported elsewhere [14], indicated that the apostrophe use of player-learners did improve.

Serious games provide a space for autonomous learning and risk taking. These can be positive and motivational attributes of learning environments. More investigation into the balance between play-learner autonomy and more direct instructional approach is required. Data analytics, transformed into different types of visualizations, can provide an objective representation of the relations between gaming elements, learning behavior and pedagogical intent. Bringing together these three domains of knowledge is important for the advancement of serious game design and for promoting genuine interdisciplinary dialogue about learning within serious games.

---

[4] The *Apostrophe Power* app is available for free download on the Apple iTunes App Store (iOS) and Google Play (Android). See http://dice.newcastle.edu.au/resources.html.

References

1. Barab, S.A., Gresalfi, M., Ingram-Goble, A.: Transformational play using games to position person, content, and context. Educ. Res. **39**(7), 525–536 (2010)
2. Browne, K., Anand, C., Gosse, E.: Gamification and serious game approaches for adult literacy tablet software. Entertainment Comput. **5**(3), 135–146 (2014)
3. Chung, G.K.W.K.: Guidelines for the design and implementation of game telemetry for serious games analytics. In: Loh, C.S., Sheng, Y., Ifenthaler, D. (eds.) Serious Games Analytics, pp. 59–79. Springer, Switzerland (2015)
4. DiCerbo, K., Bertling, M., Stephenson, S., Jia, Y., Mislevy, R., Bauer, M., Jackson, G.T.: An application of exploratory data analysis in the development of game-based assessments. In: Loh, C.S., Sheng, Y., Ifenthaler, D. (eds.) Serious Games Analytics, pp. 319–342. Springer, Switzerland (2015)
5. Gordon, C.S., Parkinson, M.J., Parsons, J., Bromfield, A., Duffy, J.: Uniqueness and reference immutability for safe parallelism. SIGPLAN Not. **47**(10), 21–40 (2012)
6. Grace, A., Kemp, N., Martin, F.H., Parrila, R.: Undergraduates' attitudes to text messaging language use and intrusions of textisms into formal writing. New Media Soc. **17**(5), 792–809 (2015)
7. Gunter, G., Kenny, R., Vick, E.: A case for a formal design paradigm for serious games. J. Int. Digital Media Arts Assoc. **3**(1), 1–19 (2006)
8. Harpstead, E., MacLellan, C.J., Aleven, V., Myers, B.A.: Replay analysis in open-ended educational games. In: Loh, C.S., Sheng, Y., Ifenthaler, D. (eds.) Serious Games Analytics, pp. 381–399. Springer, Switzerland (2015)
9. Liu, M., Kang, J., Lee, J., Winzeler, E., Liu, S.: Examining through visualization what tools learners access as they play a serious game for middle school science. In: Loh, C.S., Sheng, Y., Ifenthaler, D. (eds.) Serious Games Analytics, pp. 181–208. Springer, Switzerland (2015)
10. Loh, C.S., Sheng, Y.: Measuring expert performance for serious games analytics: from data to insights. In: Loh, C.S., Sheng, Y., Ifenthaler, D. (eds.) serious Games Analytics, pp. 101–134. Springer, Switzerland (2015)
11. Loh, C.S., Sheng, Y., Ifenthaler, D.: Serious games analytics: Theoretical framework. In: Loh, C.S., Sheng, Y., Ifenthaler, D. (eds.) Serious Games Analytics, pp. 3–29. Springer, Switzerland (2015)
12. Lukač, M.: Apostrophe(')s, who needs them? Engl. Today **30**(3), 3–4 (2014)
13. Smith, S.P., Blackmore, K., Nesbitt, K.: A meta-analysis of data collection in serious games research. In: Loh, C.S., Sheng, Y., Ifenthaler, D. (eds.) Serious Games Analytics, pp. 31–55. Springer, Switzerland (2015)
14. Smith, S.P., Hickmott, D., Bille, R., Burd, E., Southgate, E., Stephens, L.: Improving undergraduate soft skills using m-learning and serious games. In: 2015 IEEE International Conference on Teaching, Assessment, and Learning for Engineering (TALE 2015), pp. 230–235. IEEE Press, New York (2015)
15. Snow, E., Allen, L., McNamara, D.: The dynamical analysis of log data within educational games. In: Loh, C.S., Sheng, Y., Ifenthaler, D. (eds.) Serious Games Analytics, pp. 81–100. Springer, Switzerland (2015)
16. Wallner, G., Kriglstein, S.: Comparative visualization of player behavior for serious game analytics. In: Loh, C.S., Sheng, Y., Ifenthaler, D. (eds.) Serious Games Analytics, pp. 159–179. Springer, Switzerland (2015)