Gamification and serious game approaches for adult literacy tablet software

Kevin Browne[1]

Department of Computing and Software, ITB-229
McMaster University
1280 Main Street West
Hamilton, Ontario L8S 4K1
brownek@mcmaster.ca

Christopher Anand

Department of Computing and Software, ITB-202
McMaster University
1280 Main Street West
Hamilton, Ontario L8S 4K1
anandc@mcmaster.ca

Elizabeth Gosse

Brant Skills Centre, 225 Fairview Dr
Brantford, Ontario N3R 7E3
liz@brantskillscentre.org

Abstract

In this paper, we overview the design of tablet apps we designed and built to teach literacy to adults, and present the results and conclusions derived from experiments performed with target users. Low adult literacy is a significant problem with a high economic cost both for the individuals and for society. Programs created to address low adult literacy face access and engagement barriers that tablet software may be able to help overcome. We designed three tablet apps, using two contrasting approaches of incorporating game-design elements to engage the users. We tested the apps with participants from the Brant Skills Centre, a non-profit organization that offers adult literacy programs in Brantford, Ontario. Though participants were divided on whether they preferred the apps to more traditional instruction, most participants preferred using the apps in addition to more traditional instruction. Based on this we conclude that gamification and serious game design approaches were effective at increasing learner engagement, and we propose a direction for future research.

Keywords:
gamification, serious game, tablet user experience, user experience, literacy, adult literacy

1. Introduction

The high prevalence of low literacy skills in adults has serious negative consequences for society and the individuals affected. Literacy is frequently measured on a scale of 1 to 5, with level 3 being equivalent to high school completion and considered by experts as the minimum required for coping with the increasing demands of the information economy[1]. The International Adult Literacy and Skill Survey found that 48% of the Canadian population over 16, representing 12 million people, have literacy skills below this level[2].

The employment, economic and health impacts of low literacy skills have been extensively documented by international and national organizations[3, 4]. We highlight a few statistics here to illustrate the scope of the problem. People with low literacy skills are more likely to have lower rates of employment and to work

[1]Corresponding author

in low-skill jobs; roughly 57% of Canadians age 16-65 at a level 1 literacy level were employed compared to about 70% at level 2 and more than 80% of those at levels 4-5[2]. In the USA, a 2005 report by the National Assessment of Adult Literacy found that 93 million people lack the literacy skills required to complete the education or job training required by current and future jobs[5]. With employment growth polarizing into high-skill, high-wage jobs and low-skill, low-wage jobs due to automation of routine tasks by computer software and off-shoring of middle-skill jobs[6], the employment prospects of low-literacy workers will likely only become worse in the future. In the USA, an estimated $106 to $238 billion dollars in annual health-care costs can be attributed to poor health knowledge and behaviour resulting from low literacy[7, 8].

The total potential enrollment and drop-out rates of programs designed to address low literacy are disappointing. Less than half of those who reach out to a literacy organization register for a program, and of those 30% will drop out[9]. Under 10% of Canadians who could benefit from a literacy program enroll[4], with job, money problems, lack of childcare and transportation cited as reasons.

Tablet computers have recently exploded in popularity, reaching 116 million in sales worldwide in 2012 and are projected to grow to 468 million by 2017[10]. These devices provide an opportunity for novel and disruptive approaches to the problem of low adult literacy. Tablet software may be more cost effective than human instruction, alleviating the resource issue. A tablet can provide an interactive learning experience that doesn't require the learner to leave home, alleviating the lack of childcare and transportation issues. Tablet software may also provide a more fun and engaging method of learning that could increase adult participation in literacy education and / or lower literacy-program drop-out rates. Tablet devices offer a direct-control user interface (touchscreen) that could be easier to use, particularly for low-literacy users, than traditional PC or laptop home computers with keyboards and indirect pointing devices (mice, touchpads). Finally, Megalingam et al. have proposed a tablet device called EduPad to reduce rural adult illiteracy in India[11], motivated by the idea that the device could make up for a lack of qualified personnel or adequate infrastructure.

In this paper we explore using iPad software to teach literacy skills to low-literacy adults, motivated by the above situation and possible opportunity. For this we sought out existing literature on educational iPad software design and user experience, and on using gamification as a design approach due to its claim of influencing user behavior and increasing engagement.

Kayne Toukonen looked at how the features of a tablet device could facilitate learning in the form of what he termed *dynamic electronic textbooks*[12]. Toukonen pays particular attention to Brian Cambourne's eight Conditions of Learning[13] that Cambourne felt necessary for the acquisition of language, connecting each of these eight conditions to the possible software features enabled by tablets. For example, it is suggested that the learning condition *engagement* can be facilitated by using virtual worlds in tablet applications.

There have been some studies analyzing the effectiveness of tablets for education. For example, Houghton Mifflin Harcourt conducted a year-long pilot study in a middle school in Riverside, California in which an iPad application was used to increase student performance by 20% in algebra compared to peers who used textbooks[14]. In work by Sarah Henderson and Jeff Yeow, a New Zealand primary school used iPads in classrooms with students aged 5-12, and they reported that the low learning curve, high portability, and instant boot-up time of the iPads were recognized as positive aspects of using the tablets as learning tools[15]. Henderson and Yeow's findings and that of others[16] focus less on the effectiveness of the design of any particular application but more so on the aggregate results from using iPads as a learning tool; for example, the organizational challenges that using the tablets present.

The authors have previously created iPad apps for secondary and introductory post-secondary computer science education with promising results. A common design approach in these apps was having the user *learn-by-doing*; the user would manipulate representations or simulations of a problem domain. We believe this design approach is best described as experiential learning. Though experiential learning has several definitions, properties and models[17], for the purposes of our software we focus on the fundamental feature of learning through the application of concepts in an interactive setting.

Gamification is another design approach which has produced promising increases in engagement in several different contexts. Gamification can be defined as the "usage of game design elements to motivate user behaviour in non-game contexts"[18]. A well-known example would be the location-based social network Foursquare that rewards users for checking in to their current location with points and allows the user to become the "mayor" of that location. Other examples include startup company ZamZee[19] that was able to use a gamification-based app to increase physical activity in

children by 60%, and the Greater Washington Give Day who were able to use gamification to drive 2 million dollars in fundraising in one day[20].

While it's clear gamification has produced some promising results, the design approach does come with significant risk. One of the world's leading IT research and analysis companies, Gartner has warned that 80% of all gamification apps will fail to meet their objectives due to poor design[21]. There has also been little academic research into the definition of gamification itself, and whether it constitutes a truly new and distinct phenomenon. One one can also distinguish between gamification and "serious games", that is a full-fledged game for non-entertainment purposes, but this boundary is subjective.

There have been a few studies looking at educational-tablet game software. Wattanatchariya et al. developed a game called "Drop Donuts" for teaching wind and gravity concepts and report on survey results from 12 university student participants[22]. Feng Yan developed an iPad application "A Sunny Day" to educate children with autism, which features mini-game and puzzle-game elements, and tested it with autistic children and their parents[23]. Yan's work is especially valuable because of attention to application-design elements, to testing of the interface with target users, and reporting on what did and didn't work and why. For example, their analysis suggests that the software should have had clearer objectives and provided rewards more quickly after completing tasks. The paper concludes that the app offers a cost-effective therapeutic approach relative to existing methods. We were not able to find similar work examining gamification approaches to adult-literacy tablet software and consider this a significant gap in the literature that demands exploration.

In this paper, we present the design and user experience experiment results for three iPad applications developed to teach literacy concepts to low-literacy adults. The apps incorporate game-design mechanics in different ways, and we hypothesize using these apps will improve the educational experience through increased motivation, engagement, and learning. The study was done in conjunction with the Brant Skills Centre, a non-profit organization that provides programs for literacy and other essential skills for adults in Brantford, Ontario. In Section 2, we present the design of the applications and how they incorporate game mechanics. In Section 3, we discuss the design of the experiment that took place. In Section 4, we provide the the results of this study, and in Section 5 we conclude that learner engagement was improved and propose a direction for future research.

2. Application Design

Three apps were designed to teach literacy concepts to experiment participants from the Brant Skills Centre. Because the clients of the Brant Skills Centre were at varying levels of low literacy, we chose to focus on three areas that tended to be problems across these levels: homophones, punctuation (period, question mark and quotation placement), and comma placement. All of the apps incorporate some form of experiential learning by having the user *do* what the app is trying to teach them. The homophone app and punctuation app were developed with two contrasting approaches to gamification of educational software. The comma placement app was actually developed in response to experiments conducted with the homophone app, as explained in Section 4.

In a paper focusing on the definition of gamification, Deterding et al distinguish between gamification from serious games as follows, "what distinguishes gamification from regular entertainment games and serious games is that they are built with the intention of a system that includes elements from games, not a full game proper"[24]. The homophone app is intended as an educational app that includes elements of gamification to motivate user behavior, whereas the punctuation app is intended as a serious game where literacy-skill improvement is the primary goal rather than entertainment. The design of the homophone app was most influenced by Cambourne's conditions of learning, and the design of the punctuation app was most influenced by the concept of flow[25].

2.1. Homophone app

The homophone app was built to teach the differences between the following homophones:

- It's / Its
- We'll / Will
- You're / Your
- Two / Too / To
- Principle / Principal
- They're / There / Their

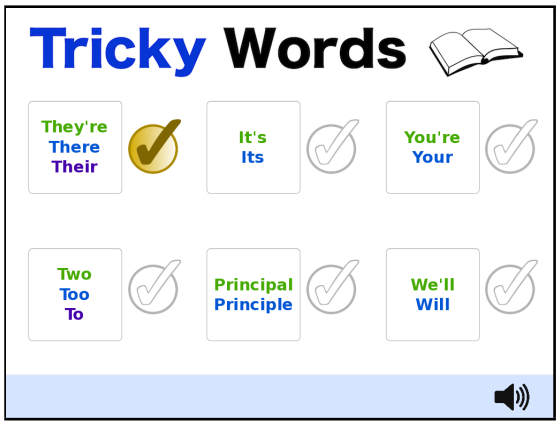

Figure 1: Title screen

Figure 2: Subscreen

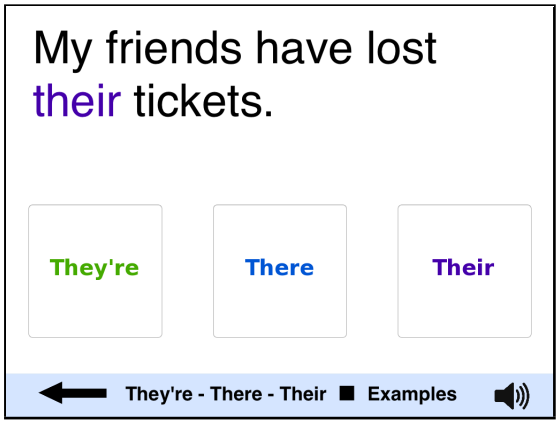

Figure 3: Examples

2.1.1. Overview

The homophone app begins with a title screen, shown in Figure 1, which presents the user with six homophone suboptions, allowing the user to choose what they would like to learn. Each option has a grayed-out check mark next to it if the user has not successfully completed the practice option provided in the submenu.

As shown in Figure 2, each homophone submenu has the same three options: lesson, examples, and practice. The homophone submenu also contains a practice score, with five smaller check marks on the left-hand side and a large check mark on the right-hand side all initially grayed out. The users can earn up to five green check marks on the left-hand side by doing the practice activity, and are awarded a gold check mark on the right-hand side (and in the title screen) when they have earned all five green check marks.

The lesson option brings the user to a new screen which gives a brief overview of the differences between the homophones containing a mixture of text and accompanying audio, illustrated with an example usage of each word.

The examples option shown in Figure 3 brings the user to a screen where they can tap a button for each homophone and are delivered an example usage of that homophone with text and audio. The user is delivered a new example each time they tap a button in the examples screen.

As shown in Figure 4, the practice option begins with a short non-interactive demonstration where a drawing of a hand drags a homophone into the blank space in a sentence and is awarded a green check mark for doing so. The practice option only does this the first time it is loaded so the user knows what is expected of them.

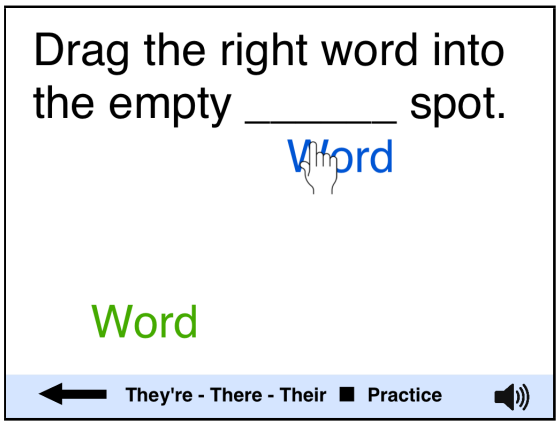

Figure 4: UI demonstration

The user is given a series of five sentences with blank

spaces for a homophone. Each sentence is presented both as text and audio. When the user drags the correct homophone into the blank space of the sentence, a chime with an approving sound is played and a green check mark moves horizontally across the screen as shown in Figure 5.

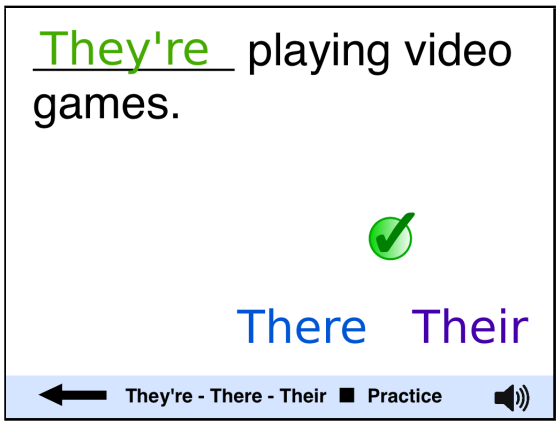

Figure 5: Right answer

When the user drags an incorrect homophone into the blank space of the sentence, a chime with a disapproving sound is played and a yellow circle with a diagonal line through it (i.e. a no symbol) moves horizontally across the screen as shown in Figure 6. The yellow no symbol was chosen because the usual red symbol would be too discouraging for participants. Though the meaning of sound effects are subjective and difficult to describe in writing, the disapproving sound for an incorrect answer was similarly sensitive in that it was more akin to a "missed shot" sound than the sort of "buzzing error" sound that you might hear on a TV game show.

Figure 6: Wrong answer

The highest score out of five that the user has achieved is presented on the associated homophone submenu. The user can try the practice option as many times as they want. A back-arrow button in the bottom left allows users to move back to the previous menu at all times, and a speaker button at the bottom right allows the user to turn on or turn off the audio.

2.1.2. Design approach

The overall design of the app is motivated by Cambourne's conditions of learning and Kayne's interpretations for how tablet software can fulfill these conditions[12]. The conditions are summarized as follows:

1. **Immersion** - learners need to be immersed and constantly saturated in that which is to be learned.
2. **Demonstration** - learners need to receive many stimulating demonstrations of desired outcomes.
3. **Engagement** - learners must be engaged in the learning process while being immersed in the learning environment and viewing demonstrations.
4. **Expectations** - learners are influenced by expectations, which are powerful shapers of behavior.
5. **Responsibility** - learners need to make their own decisions about when, how, and what "bits" to learn.
6. **Employment** - learners need time and opportunity to use and practice new learning in realistic ways.
7. **Approximation** - learners must be free to approximate desired study, as mistakes are essential for learning to occur.
8. **Response** - learners must receive relevant, appropriate, timely, nonthreatening feedback.

The app is designed to give users *responsibility* by allowing users to decide which homophones to learn (title screen) and how to learn them (through a lesson, examples and / or practice). No portion of the app has a time limit so users are free to learn when they want to learn as well. *Immersion* is facilitated by the audio that accompanies the text throughout the app. The examples feature of the app aims to fulfill the *demonstration* learning condition. *Expectations* are set out for the user by showing the grayed-out check marks when the app first launches and allowing them to earn green and gold check marks through successful practice. *Employment* of the concepts is provided by the practice feature of the app which allows users to practice their understanding in a realistic setting. The users are free to *approximate* their understanding as there is no penalty for making mistakes during the Practice portion of the app;

users can redo the Practice portion as much as they like and are only rewarded with check marks for correct answers and do not lose anything for incorrect answers. These check marks and yellow no symbols that the user receives during practice facilitate the *response* learning condition.

Finally the *engagement* condition is facilitated by the gamification elements of the app:

- **Badges** - green and gold check marks as rewards for successful practice.
- **Levels** - each homophone submenu provides a discrete objective for the user.
- **Short, medium, and long term goals** - earning an individual green check mark is a short-term goal, earning a gold check mark for a homophone submenu is a medium-term goal, and earning all of the gold check marks is a long-term goal.

Distinguishing between short-term goals with a green check mark and medium-term goals with a gold check mark creates a visual layering of goals that take increasing amounts of time to complete but that come with increasing rewards is a key feature of many games[26].

2.2. *Punctuation app*

The punctuation app was built to teach when to use a period at the end of a sentence, when to use a question mark, and where quotation marks should be used (around phrases to show what words are being said, as well as around titles of stories, songs, books and movies).

2.2.1. *Overview*

The punctuation app does not begin with any sort of menu, but begins immediately with an introductory demonstration where a period is dragged by a drawing of a hand to the end of a sentence. From this screen shown in Figure 7, the user is then left to drag the remaining period to the end of the next sentence. A horizontal bar at the top of the screen next to a drawing of a clock tells the user how long they have to drag the period to the end of the sentence; the bar will turn yellow when only 25% of the time alloted is remaining.

If the user completes this step, then as seen in Figure 8, they will begin a virtual "tour of Italy" that simultaneously attempts to teach the user the correct usage of periods, question marks, and quotation marks. Each screen presents the user with a series of sentences which require the appropriate punctuation marks to be dragged

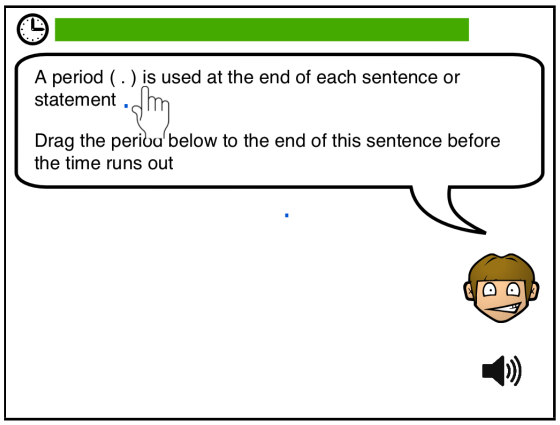

Figure 7: Punctuation app introduction

into the correct places before the timer runs out in order to move on to the next screen. The user receives an explanation about how to use each punctuation mark correctly in the first screen that the punctuation mark is used. If the user drags a punctuation mark into the correct position, an approving sound is played. If the user drags a punctuation mark into an incorrect position, it moves back to its original position below the text. If the user does not complete the task in time, the correct answer is presented to the user as an audio explanation as to why this answer is correct. The user then repeats the same screen until they drag the correct punctuation marks into the correct positions.

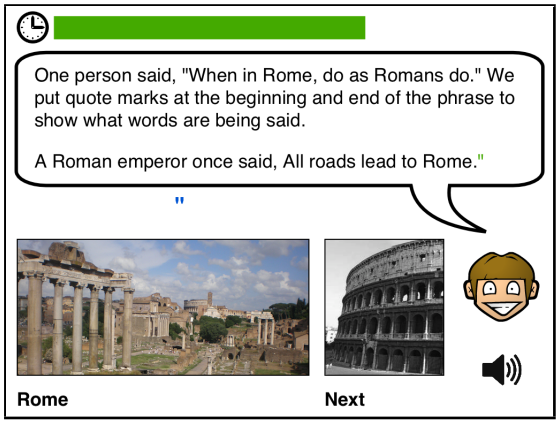

Figure 8: Punctuation app tour

The screens of the punctuation app get increasingly difficult over time, as at first the user is given only one or two punctuation marks to drag into the correct position before being given three and then finally four as shown

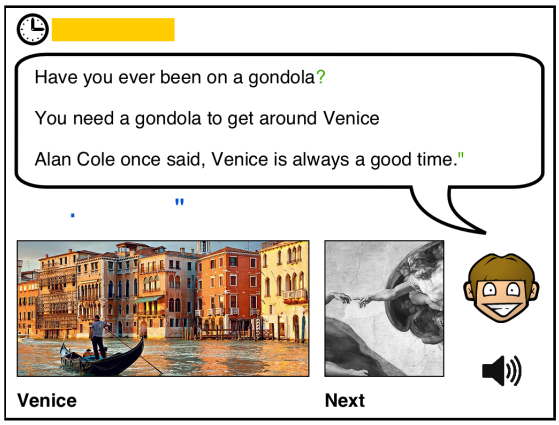

Figure 9: Increasing difficulty

in Figure 9.

There are ten screens in the punctuation app. All of the text in each screen is also presented to the user as audio. When the user completes the final screen, they are given a congratulatory goodbye by the character who has guided them through the tour in the lower right-hand side of the screen. The images of Italy that appear throughout the app were intended to provide an aesthetically pleasing reward for reaching each screen, with a grayed-out image of the next screen to provide some motivation to see what's next.

2.2.2. *Design approach*

In contrast to the "learning system with game design elements" approach of the homophone app where the user self-consciously directs what, how, and when they learn, the punctuation app was designed to have the users experience *flow* and a loss of reflective self-consciousness as they become absorbed in playing the game. Nakamura and Csíkszentmihályi[27] describe flow as a subjective experience that seamlessly unfolds from moment to moment with the following features:

- intense and focused concentration on the present moment;
- merging of action and awareness;
- loss of reflective self-consciousness;
- a feeling of control over the activity;
- distortion of temporal experience (i.e. that time has passed faster than normal);
- experience of the activity as intrinsically rewarding.

They also define the conditions for entering a state of flow to include the perception of challenges that stretch but do not exceed existing skills, as well as clear goals and immediate feedback about progress being made. As adapted from Csíkszentmihályi[25], the earlier model of flow emphasized a balance between perceived opportunities and skills as shown in Figure 10. The current model of flow shown in Figure 11 has apathy experienced when the perceived challenges and skills are below the user's average levels, and flow experienced when the challenges and skills are above the user's average levels (i.e. the stretching of existing skills).

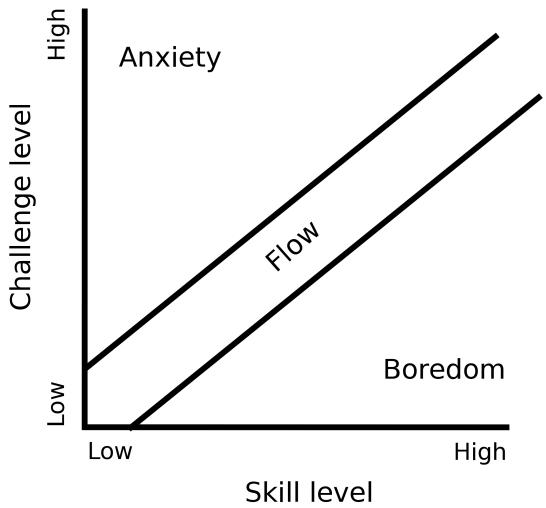

Figure 10: Earlier model of flow

The punctuation app was designed to induce a sense of flow through the clear goals and immediate feedback of moving the punctuation marks to the correct positions in the text. The timer at the top of the screen was meant to help focus the user's attention on the present moment as it would discourage them from losing focus on the task at hand. A tour of Italy was chosen as the narrative for the app because vacations and tours are intrinsically motivating activities.

Facilitating the perception that challenges stretch but do not exceed skills was done in two ways. First, any time a new concept was introduced, it is explained to the user (how the UI functions, the correct usage of periods, question marks, and quotation marks). This was done to ensure that the challenges being posed do not exceed the user's perceived skills. Second, the app gets progressively more difficult from one screen to the next by introducing more difficult problems (more text, more

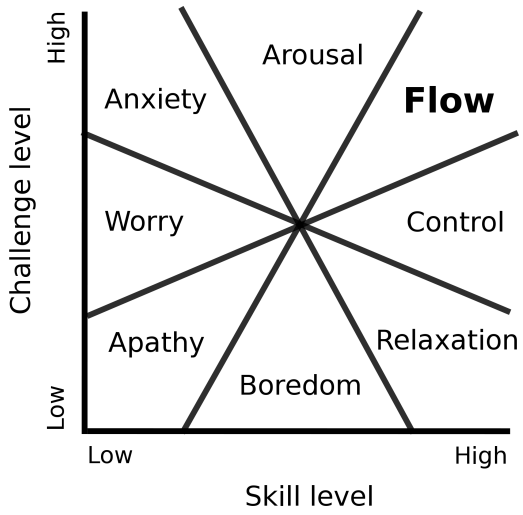

Figure 11: Csíkszentmihályi model of flow

punctuation marks to place on the text). This was done to progressively stretch and ideally expand the user's skills, like climbing a set of stairs to a better understanding of the concepts. While more advanced users could become apathetic when faced with the initial easy screens which do not stretch their skills, we try to mitigate this by allowing advanced users to quickly reach the more difficult screens.

Other game features besides flow incorporated into the design of the app include the *narrative* of the tour of Italy which involves connected events with a beginning and an end for the user to progress through. And just like stops on a tour bus, each stop has a maximum duration made visually apparent by time bar.

2.3. *Comma app*

The comma app resembles the practice feature of the punctuation app, but with a few differences. The app begins with a demonstration where an image of a hand drags a comma to the correct position in a sentence. The user is then given a cycle of five sentences where they must drag the comma to the correct position. If the user drags the comma to the correct position they receive an approving sound and a green check mark as shown in Figure 12.

If the user drags the comma to an incorrect position, then, as shown in Figure 13, the correct answer is explained using audio. From here the user moves on to the next of the five sentences which has not been completed correctly.

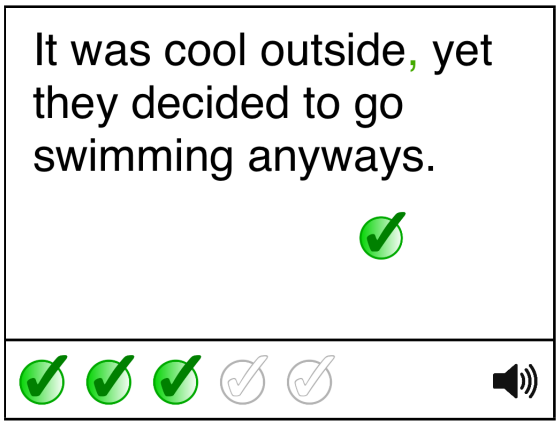

Figure 12: Comma app right answer

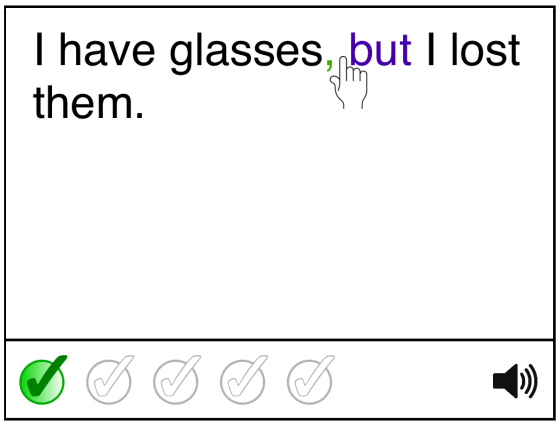

Figure 13: Correct answer demonstration

The five sentences remain in the cycle until all of them have been completed correctly. A series of five grayed-out check marks in the bottom left-hand corner of the screen turn into green check marks as the user correctly completes each sentence. When the user completes all five sentences correctly, they are awarded a gold check mark and the app is complete.

2.3.1. *Design approach*

Experiments conducted with the homophone app inspired the creation and subsequent test of the comma app, which is essentially a modified version of the practice feature of the homophone app, and we explain the reception of the app in Section 4.

3. Experiment Design

Clients of the Brant Skills Centre were invited to participate in a study about teaching literacy with iPad applications. Staff spoke to them directly in conversations

or during class. Clients who chose to participate were given consent forms to sign to confirm their participation and make them aware of the study details. We made sure that the consent forms for this study were written in plain language, and that if the client was unable to read the consent form unassisted or had any questions, that the details were also be explained to them verbally to ensure that we had their consent.

The sessions took place in a classroom at the Brant Skills Centre and were conducted by Kevin Browne and Elizabeth Gosse. Participants received a $10 Tim Hortons gift card for each experiment session that they participated in to compensate them for their time and motivate their participation in the study. The study was approved by the McMaster Research Ethics Board.

3.1. Experiment session protocol

The following protocol was followed with each experiment participant. The protocol refers to the pre-experiment questionnaire in Section 3.4, the user experience survey in Section 3.5, and the post-experiment questionnaire in Section 3.6.

Four one-hour experiment sessions were conducted, two sessions in one week and two sessions in the following week. Participants could only attend one session per week, and some participants did not attend a session in both weeks. In the first week's session, the homophone app was tested, and in the second week's session, the punctuation app was tested followed by the comma app. The following procedure was used during each session for the homophone-app and punctuation-app experiments:

1. The participants were told the goal of the experiment.
2. The rest of the experiment procedure was outlined for the participants.
3. The participants completed a paper copy of the pre-experiment questionnaire.
4. The participants completed the first quiz sheet.
5. The participants were instructed using either traditional methods or the iPad app.
6. The participants completed the second quiz sheet, and the user experience survey if the iPad app had just been used.
7. The participants were instructed using either traditional methods or the iPad app.
8. The participants completed the third quiz sheet, and the user experience survey if the iPad app had just been used.
9. The participants completed the post-experiment questionnaire.

Each week, both presentation orders were used in one session (iPad app followed by traditional instruction versus traditional instruction followed by iPad app). The time spent on each method was roughly equivalent, about 15 or so minutes, or until the participants felt like going ahead and doing the next practice quiz.

Traditional instruction consisted of Brant Skills Centre instructor Gosse teaching the concepts using a lecture and whiteboard, as well as pen-and-paper practice questions (different from the quiz sheets used in the above procedure). Gosse's lecture method was highly interactive and tailored to the unique learning styles of the participants, as discussed in Section 4.

During the second week's session, after following the above procedure, we also did a brief experiment of the comma app. As is discussed further in Section 4, the comma app was developed quickly in response to the results of the first week's experiment that indicated a combination of traditional instruction and the iPad software may be most helpful. In the first session of the second week, the participants were given a brief lecture on comma placement, then practiced their understanding with pen and paper practice sheets, followed by another brief lecture on comma placement, followed by practice with the comma app. In the second session of the second week, the order of the lecture and accompanying practice method was reversed. In both instances, the participants completed the comma-experiment questionnaire found in Section 3.7.

The first week's sessions had a combined 14 participants (8 in the first session and 6 in the second), and the second week's sessions had a combined 13 participants (7 in the first session and 6 in the second).

3.2. Quantitative observations

Quantitative observations were recorded using the quiz sheets to test the participants throughout the homophone-app and punctuation- app experiments. The quiz sheets consisted of one page of relevant questions. In the case of the homophone experiment, the participants were asked to fill in the blank word of a sentence with the correct word. In the case of the punctuation experiment, the participants were asked to fill in the missing punctuation of a sentence. The same quiz was given three times in an attempt to assess if learning was taking place, and how much, over the course of the experiments. The homophone quiz had a maximum score of 19, and the punctuation quiz had a maximum score of 10.

3.3. Qualitative observations

After the session was over, a casual verbal discussion with the participants as a group and participants individually was used for further insights into the effectiveness of the software. Observations from these discussions were recorded in writing. Verbal expressions, reactions, and comments made by the participants during the study were also be recorded in writing as study data. Qualitative observations as to how users perceived each app were also recorded with the user experience survey in Section 3.5.

3.4. Pre-experiment questionnaire

The following information was gathered with the pre-experiment questionnaire:

- Gender (Male/Female)
- Age
- Handedness (Right/Left)

The participants were also asked to rate their reading ability from 1 to 5, if 1 is "not well at all" and 5 is "I can read perfectly well", and asked to rate their ability to use the iPad from 1 to 5, if 1 is "not well at all" and 5 is "I can use the iPad perfectly well".

3.5. User experience survey

The participants were asked to rate how much they agree (Likert scale) with the following statements:

- **S1** The app was easy to use.
- **S2** It was easy to learn how to use this app.
- **S3** I enjoyed using this app.
- **S4** The iPad was comfortable to hold while using the app.
- **S5** The app helped me to learn _____.
- **S6** I found the app to be useful.
- **S7** I would tell other people to use this app.
- **S8** The touchscreen finger gestures required to use the app felt natural.
- **S9** I liked the app's graphics.
- **S10** I liked the app's voices / sound.

The participants could choose from: strongly disagree, disagree, neutral, agree, and strongly agree. For analysis purposes, these descriptions were assigned numeric values 1-5 from strongly disagree to strongly agree. An overall user experience score from 1-5 can also be computed by averaging the assigned numeric values. In the case of statement S5, the blank space was replaced with the concept being taught in the session. This and other wordings of the user experience survey statements were done under the advisement of Brant Skills Centre staff.

3.6. Post-experiment questionnaire

The following questions were asked on the post-experiment questionnaire.

1. Did you prefer learning using the iPad app or learning from the instructor's lesson? (check one)
2. In the future should _____ be taught only using the iPad app, only using the instructor's lesson, or both? (check one)

The blank space was again replaced with the concept being taught in the session.

3.7. Comma-experiment questionnaire

The following questions were asked on the comma-experiment questionnaire.

1. I liked using the app to do practice exercises during the lesson.
2. In the future, I would prefer doing practice exercises during lessons using the iPad app, only using the instructor's lesson, or either? (check one)

The first question allowed participants to choose from: strongly disagree, disagree, neutral, agree, and strongly agree. Again for analysis purposes, these descriptions were assigned numeric values 1-5 from strongly disagree to strongly agree.

4. Results and Discussion

The first week of sessions where the homophone experiment took place had 14 participants total, 8 males and 6 females, 2 left-handed and 12 right-handed, with ages ranging from 21 to 63 and an average age of 41 (standard deviation of 13.5). The second week of sessions where the punctuation experiment and comma experiment took place had 13 participants total, 7 males and 6 females, 3 left-handed and 10 right-handed, with ages ranging from 21 to 54 and an average age of 42 (standard deviation of 11).

The participants' ratings of their reading ability and ability to use an iPad are presented in Figure 14. The participants' ratings of their ability to use an iPad are noticeably higher in the second week of sessions, which may indicate increased confidence in using the iPads after doing so during the first week's sessions.

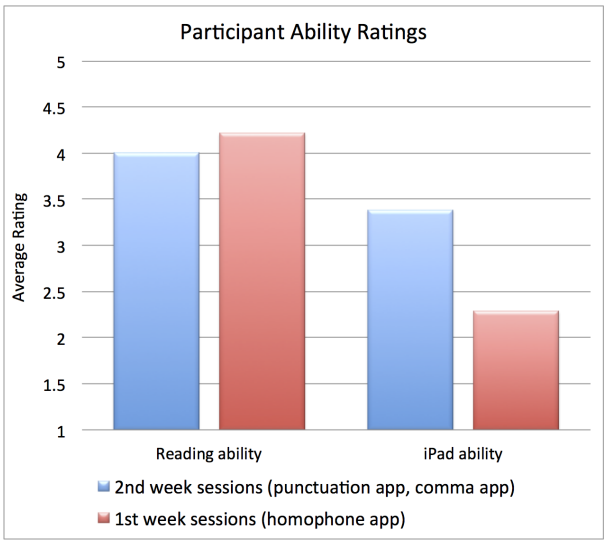

Figure 14: Participant ability ratings

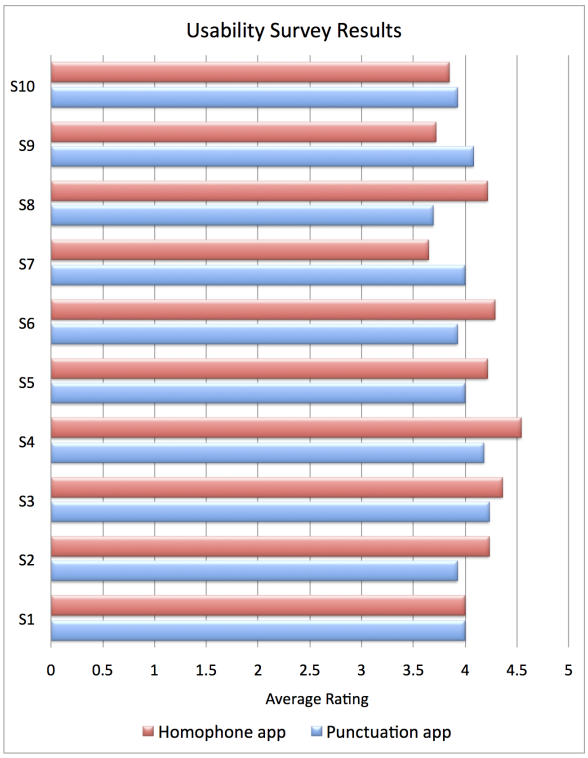

Figure 15: User experience survey results

4.1. Homophone experiment

The homophone app received overall positive user experience feedback as shown in the survey results in Figure 15. The lesson and example features of the app were used by the participants and by all accounts were well received, but it was definitely the practice feature of the app that participants used most.

The layered goals provided by the green and gold check marks were very effective at motivating user behavior as participants genuinely became emotionally involved in obtaining the correct answers. Exclamations like "Oh come on!" and "This is fun!" were recorded during the sessions, as well as laughter. Virtually every participant continued to work through the practice feature of each homophone submenu in the app until they had earned a gold check mark.

Though the homophone app and experiment were not designed with a social aspect in mind, by virtue of the fact that participants were sitting around the same large table, the participants' interactions with one another played a role. Several of the participants playfully competed against one another to complete the app by obtaining the gold check marks. Several participants also tilted their iPads so that they could see their neighbors screens and help each other out. Every participant discussed *some* aspect of the app while using the app with either the other participants, Gosse, or Browne.

One negative observation was initial frustration experienced by several users at what to do with the app from the beginning title screen. They felt confused by the amount of choice that they were provided, without any direction as to what to do. The confusion was momentary as we either explained what to do or the participants eventually experimented with the app by tapping one of the six buttons and entering a homophone submenu. Participants told us that this confusion could have been alleviated with audio providing direction.

Another negative observation was one user's frustration that the practice feature would force them to redo sentences that they had already got right during a previous attempt. They were stuck on the last of the five practice questions and were having to repeat all of the previous questions to keep trying to earn the gold check mark.

The results of the three homophone quizzes are shown in Figure 16, with separated average scores from the two sessions and the combined average. Overall in

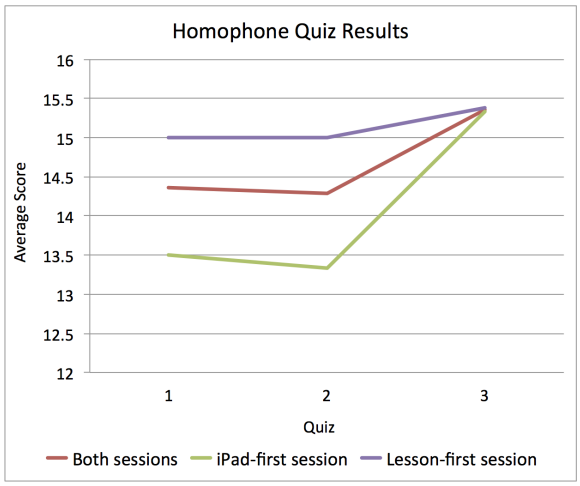

Figure 16: Homophone quiz results

both sessions, the average quiz score slightly improved after both instructional methods.

The post-experiment questionnaire results regarding the preferred learning method and suggested future learning method(s) are found in Table 1 and Table 2. That the majority of participants suggested that in the future homophones be taught using both traditional methods and the app fits very well with observations and post-experiment discussions with participants.

Instruction method	**Homophone sessions**	**Punctuation sessions**
iPad	7	7
Lesson	7	6

Table 1: Preferred learning method

Instruction method	**Homophone sessions**	**Punctuation sessions**
iPad	0	0
Lesson	2	2
Both iPad & lesson	12	11

Table 2: Suggested future learning method

Though perhaps it may seem an obvious point, co-researcher Browne found it remarkable observing just how effective a human instructor can be at teaching literacy. Many participants had unique challenges to learning, whether it was their hearing ability, ability to stay focused, difficulty in grasping a concept, etc. As an instructor, Gosse was able to customize her approach to each individual participant's learning needs, from moment to moment as the situation evolved, in an emotionally sensitive manner. These abilities are very difficult for any piece of software to ever fully duplicate. The importance of these abilities during instruction may be even more important in a domain such as literacy that involves natural language and human-to-human communication by nature of the subject.

Discussions with the participants revealed that while many preferred to be taught the concepts via a traditional lesson, many preferred to practice and reinforce their understanding of the concepts with the iPad app. This view likely explains why the majority of participants suggested using both instructional methods in the future.

The discussion also revealed a strong consensus that after the users suffered an incorrect answer during the practice feature, the app should explain the correct answer to the user rather than displaying symbols that indicate an incorrect answer as the participants found this discouraging.

The results of this experiment prompted us to create the comma app, which as discussed in Section 2.3 follows a simple design similar to the practice feature of the homophone app. The comma app explains correct answers and does not force users to complete answers to already correct questions by instead cycling through the questions until they are all answered correctly.

The results for each individual participant are provided in Table 3, broken down into the session that started with the traditional method and the session that started with the iPad app. We note that of the six participants that rated their ability with the iPad as a '1', only one of these participants preferred the iPad app to the traditional lesson. The two participants who suggested that only the traditional lesson be used in the future also rated their ability with the iPad as a '1'. In contrast the three participants that rated their ability with the iPad as '4' or higher all preferred the iPad to the traditional lesson. Though more participants would be required to make stronger claims, this data provides some indication that a participant's ability to use the iPad may impact their instructional preference.

Traditonal method first session								
P	**RA**	**iA**	**Q1**	**Q2**	**Q3**	**US**	**PM**	**SFM**
P1	4	1	17	17	17	4.1	Lesson	Both
P2	3	1	7	8	11	3.1	Lesson	Lesson
P3	5	1	16	17	17	3.3	Lesson	Lesson
P4	4	1	14	16	18	3.9	Lesson	Both
P5	4	1	16	15	14	4.7	iPad	Both
P6	4	3	17	18	17	4.6	Lesson	Both
P7	4	4	19	19	19	4.2	iPad	Both
P8	5	5	14	10	10	5.0	iPad	Both
iPad method first session								
P	**RA**	**iA**	**Q1**	**Q2**	**Q3**	**US**	**PM**	**SFM**
P9	2	1	4	1	5	1.7	Lesson	Both
P10	4	2	15	14	16	4.7	iPad	Both
P11	5	2	13	8	15	4.5	iPad	Both
P12	5	3	16	19	18	4.6	iPad	Both
P13	5	2	15	19	19	4.1	Lesson	Both
P14	5	5	18	19	19	4.7	iPad	Both

Table 3: Homophone experiment participant results - P = Participant, RA = Reading ability, iA = iPad ability, Q# = Quiz #, US = User experience score, PM = Preferred method, SFM = Suggested future method

Regarding the relationship between quiz performance and instructional method preference, we observe that high performing participants P1, P3, P4, P6 and P13 preferred the lesson, and high performing participants P7, P12 and P14 preferred the iPad app. This may indicate that quiz performance was not related to instructional method preference, however we also observe that lower performing participants P2 and P9 were also responsible for the lowest user experience score results (and both preferred the traditional lesson).

4.2. Punctuation experiment

The punctuation app also received overall positive user experience feedback as shown in the survey results in Figure 15. In stark contrast to how the homophone app was discussed while it was being used, very little was spoken by participants while they were using the punctuation app. A couple of participants had difficulties that will be discussed, and they vocalized those concerns while using the app, but the remaining participants largely remained silent until they had completed the app. This would be consistent with the participants experiencing a sensation of flow as the app intended. The fact that the app explained the correct answer if the user failed to pass a screen was well received, as was the clear direction as to what the app expected the user to do next.

Participants P1 and P6 (see Table 4) in particular had a very difficult time completing the app, for reasons that could have been alleviated had the design been different. In both cases, participants had rated their ability to use the iPad as "1" and were struggling to understand the concepts themselves. One of the participants had particular difficulty dragging the punctuation marks across the screen. The participant was lifting their finger from the screen while dragging the punctuation mark, and it would move back into its original position. Each time the user failed to pass a screen, they were given the exact same amount of time as before to pass the screen on the next try. Even when these participants knew the correct answer, they simply couldn't move the punctuation marks quickly enough. They would then become more frustrated when having to sit through the same explanation of the correct answer, and have to go through the process again.

Giving participants more time on "retries" and giving them different (but perhaps equally difficult) problems to work through could likely have prevented this frustration. Giving users more time or different problems could have been accomplished statically with different optional levels of difficulty at the start of the app, or dynamically as the user experienced the app. Also, instead of having the punctuation marks snap back to their original position after being placed in an incorrect position, they could have remained in the new but incorrect position, reducing this frustration as well.

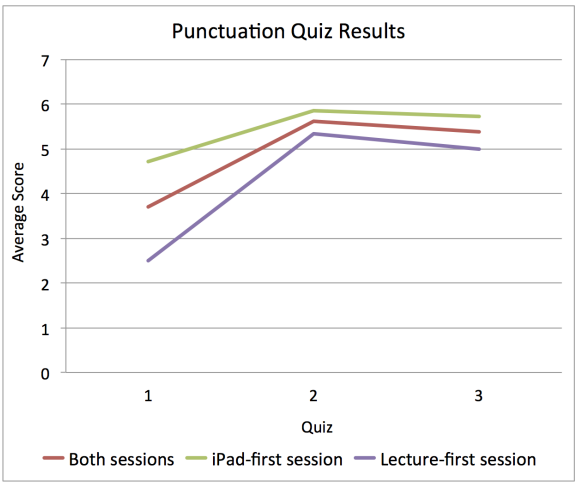

Figure 17: Punctuation quiz results

Though less evident than the homophone app, again the relative dynamism of the instructor relative to the app was very apparent. Whether learning during the les-

son or with the app, when participants failed to grasp something or get the right answer, they were understandably frustrated. In most cases by explaining the correct answer to the participant, the app was able to quell this frustration. But as noted above, sometimes this wasn't enough. Whereas the instructor was almost immediately aware of any frustration and able to deal with it appropriately, in a manner tailored to the participant.

During the punctuation sessions, the quiz results as presented in Figure 17 show an initial increase in score after the first instructional method (either traditional lesson or iPad) and a decrease after the second method. Perhaps some participants became bored after learning the same skills twice in a row.

As with the homophone app, the post-experiment questionnaire found in Table 1 and Table 2 suggest that while about half of the participants preferred each method, there was near consensus that both methods should be used in the future.

Traditonal method first session								
P	**RA**	**iA**	**Q1**	**Q2**	**Q3**	**US**	**PM**	**SFM**
P1	3	1	1	2	1	3.0	iPad	Both
P2	5	4	6	7	7	4.0	iPad	Both
P3	4	4	0	8	5	4.8	iPad	Both
P4	3	5	2	4	5	4.0	iPad	Both
P5	5	5	6	10	10	4.1	Lesson	Both
P6	1	1	0	1	2	2.1	Lesson	Lesson
iPad method first session								
P	**RA**	**iA**	**Q1**	**Q2**	**Q3**	**US**	**PM**	**SFM**
P7	4	3	6	5	5	1.7	Lesson	Both
P8	5	5	2	2	2	4.7	Lesson	Both
P9	4	4	6	10	10	4.5	iPad	Both
P10	5	5	3	3	3	4.6	Lesson	Lesson
P11	4	1	3	4	3	4.1	iPad	Both
P12	5	1	9	10	10	4.7	Lesson	Both
P13	4	5	4	7	7	4.7	iPad	Both

Table 4: Punctuation experiment participant results - P = Participant, RA = Reading ability, iA = iPad ability, Q# = Quiz #, US = User experience score, PM = Preferred method, SFM = Suggested future method

The results for each individual participant are provided in Table 4, again broken down into the session that started with the traditional method and the session that started with the iPad app. Note that these participant numbers do not represent the same participants as the homophone app sessions presented in Table 3.

Two of the four participants who rated their ability with the iPad as a '1' preferred the traditional lesson to the iPad app. Five of the eight participants who rated their ability with the iPad as '4' or higher preferred the iPad app. Again this shows further studies with more participants would be required to make strong arguments about the relationship between the ability of the participant to use the iPad and instructional method preference. The same is true of establishing relationships between quiz performance and instructional method performance, and other finer grain relationships of interest.

4.3. Comma app

The level of agreement with the first question of the comma-experiment questionnaire, "I liked using the app to do practice exercises during the lesson", was 4.4, indicating the app was well received. Of the 13 participants in the comma-app experiment, 5 reported that in the future they would prefer to do practice questions during lessons using the app, 6 said they would prefer either iPad or pen-and-paper practice questions, and 2 said they would prefer pen-and-paper practice questions.

Participants P1 and P6 (see Table 5) that said they would prefer pen and paper were the same two participants that rated their iPad ability as "1" and struggled with the homophone app. Though not experiencing as much frustration with the homophone app, both participants were again frustrated with the comma app. Again difficulty in moving the object across the screen due to lifting the finger off the screen was an issue, as well as frustration due to having the same question repeated again and again after answering it incorrectly.

Other than these instances the app was well received, with the cycling of questions and explanation of the correct answers being particularly well received.

The results for each individual participant are provided in Table 5, again broken down into the session that started with the traditional method and the session that started with the iPad app. Note that these participant numbers are the same participants as the punctuation app sessions presented in Table 4, as the brief experiment with the comma app was done as an addendum to the punctuation experiment sessions.

Traditonal method first session				
P	**RA**	**iA**	**SQ**	**FPM**
P1	3	1	3	Paper
P2	5	4	5	Either
P3	4	4	5	iPad
P4	3	5	4	Either
P5	5	5	4	Either
P6	1	1	3	Paper
iPad method first session				
P	**RA**	**iA**	**SQ**	**FPM**
P7	4	3	4	iPad
P8	5	5	5	iPad
P9	4	4	5	Either
P10	5	5	5	iPad
P11	4	1	4	Either
P12	5	1	5	Either
P13	4	5	5	iPad

Table 5: Comma experiment participant results - P = Participant, RA = Reading ability, iA = iPad ability, SQ = Survey likert question ('liked using the app'), FPM = Future preferred method

5. Conclusion

We conclude that integrating educational software incorporating game design elements into adult literacy education does increase learner engagement, on the basis of the participants' recommendation that the apps be used as part of future instruction in addition to traditional instruction. In addition, the participants' social activity and verbal expressions suggesting emotional involvement while using the homophone app, and the participants' relative silence while focused on using the punctuation app, are both evidence of contrasting forms of user engagement.

The gamification design approach of the homophone app facilitated an externally focused form of engagement. The layered goals of the app could be achieved at the participant's leisure as there was no time constraint. This allowed for participants to express their emotional state to one another, to compete with one another, and to help one another.

The serious game design approach of the punctuation app facilitated an internally focused form of engagement. The time limit kept participants focused on the challenge at hand, and the increasing difficulty of the challenges discouraged participants from becoming bored.

The external focus facilitated by the gamification design approach is consistent with the idea that gamification is about providing external rewards for the intrinsic motivations of users. The internally focused form of engagement facilitated by the serious game design approach is consistent with the idea that experiencing flow involves the loss of reflective self-consciousness as users become absorbed in the activity at hand. In either case participants were focused and engaged in the educational activity. We suggest both forms of engagement may be effective, depending on the desired approach and goals of the instruction.

The number of participants in this study does not allow us to generalize the more nuanced observations and results to the entire low-literacy adult population. For example, although it is encouraging that performance generally went up by the end of the homophone and punctuation sessions, the quiz results don't provide a strong enough signal to make broad claims about the educational effectiveness of one method or the other. For this, a further study with more participants or taking place over a larger amount of time would be necessary.

Our work and results provide re-confirmation of several user interface design concepts outlined previously in the literature (albeit in a new problem domain - adult literacy educational software):

- Use different levels of rewards (short, medium, and long term) to motivate behavior[28].
- Use time limits and gradually increasing challenge difficulty (both in time allotted and conceptual difficulty) to encourage the user to experience flow during challenge(s)[29].
- Immediately present a new challenge after the user has completed their attempts at the current challenge to encourage the user to experience flow[30, 31].
- Don't present a user with the same challenge if they have already failed the challenge twice[32].

This work was motivated by the high economic and personal costs of low adult literacy and the potential for tablet educational software based on gamification and serious game design approaches to offer novel solutions that may increase engagement, improve cost effectiveness, be easier to use, and / or have the ability to be used remotely at the user's home.

The majority of participants were able to use the software without experiencing too much frustration, and it appears that design improvements could have prevented the frustration that was experienced. This indicates that ease of use is indeed a positive factor for using tablets for adult literacy education.

While a majority of participants suggested the apps be used when the concepts are taught in the future, this was in addition to a traditional lesson. If a human instructor is still required, this limits the potential to improve cost effectiveness, or the ability of the user to benefit from being able to learn at home.

Regarding cost effectiveness, an instructor naturally has a limited amount of time that can be allotted to some mix of lessons, one-on-one tutoring, marking practice questions, etc. The positive reception of the homophone app's practice feature and the comma app suggests that instructors could automate at least some of the time spent marking practice questions using tablet apps. Using educational apps incorporating gamification for drill and practice would correspond with the historical role of video games in the classroom[33].

Landers and Callan[34] use a hammer and nail analogy to suggest educators pick a game as an instructional method only *after* they have specified what concepts are to be instructed. If adult literacy instructors could easily pick through suites of effectively designed apps made to drill and practice specific literacy concepts, they could use them to replace this portion of their time, increasing the cost effectiveness of literacy instruction.

The fact that participants suggested both a human instructor and an iPad app be used in future instruction of the concepts doesn't mean the concepts couldn't be learned from home via the apps if necessary. But looking towards future work in this area, a holy grail for literacy educational software would be completely automated instruction of all concepts done at such a high quality that learners *prefer* the apps to a human. This may very well be impossible, but we suggest a design approach that may aid in moving towards this goal.

In our experiments, it was the dynamism of the instructor that appeared to make them so effective at teaching the concepts themselves. In-the-moment adjustments based on participant reactions, learning tailored to individual needs, and emotional sensitivity all played a clear role in making the instructor effective in this way.

Our apps came closest to this sort of dynamism when they themselves were dynamic. The stair-climbing flow-based design of the punctuation app allowed for advanced individuals to proceed quickly through the app, with jumps in challenge difficulty that were accommodating enough for most participants. There was also a strong positive reception of other "dynamic" features like explanations of correct answers after suffering an incorrect answer.

We suggest that a high degree of dynamic behavior in future adult-literacy educational apps will have a positive effect on their reception and ability to replace instructors in roles beyond drill and practice. In particular, we put forward a concept of *dynamic flow inducement* for exploration.

The concept of flow offers promise in educational software beyond drill and practice because by nature it involves the stretching of skills to meet new challenges, which can then be stretched further and further, to build up a more sophisticated understanding over time. The problem with using flow in this "stair-climbing to understanding" strategy is that users all start off from different levels of understanding and abilities, and will proceed in their growth of understanding at different rates. An app could account for these issues and sustain an experience of flow by dynamically adjusting the difficulty of the challenges, the way the challenges are presented, the time requirements, etc. We propose that this concept be explored with more extensive experiments designed to test the effectiveness of strategies for maintaining an experience of flow while dynamically adjusting the challenges presented to the user.

[1] R. Desjardins, T. Murray, A. Tuijnman, Learning a living: First results of the Adult Literacy and Life Skills survey, OECD, 2005.

[2] L. Barr-Telford, F. Nault, J. Pignal, Building on our competencies: Canadian results of the international adult literacy and skills survey, Statistics Canada. Available at: www. statcan. ca/bsolc/english/bsolc (2005).

[3] R. Higher, America: Overcoming the crisis in the us workforce, 2008.

[4] B. Literacy, Who Wants to Learn?: Patterns of Participation in Canadian Literacy and Upgrading Programs: Summary of Results from the National Study Conducted for ABC Canada in Partnership with Literacy BC., ABC Canada, 2001.

[5] M. Kutner, E. Greenberg, J. Baer, National assessment of adult literacy (naal): A first look at the literacy of americas adults in the 21st century (report no. nces 2006–470), Washington, DC: National Center for Education Statistics, US Department of Education (2005).

[6] D. Autor, et al., The polarization of job opportunities in the us labor market: Implications for employment and earnings, Center for American Progress and The Hamilton Project (2010).

[7] J. Vernon, A. Trujillo, S. Rosenbaum, D. B, Low health literacy: implications for national health policy., University of Connecticut. National Bureau of Economic Research, Storrs, CT (2007).

[8] H. S. Yin, M. Johnson, A. L. Mendelsohn, M. A. Abrams, L. M. Sanders, B. P. Dreyer, The health literacy of parents in the united states: a nationally representative study, Pediatrics 124 (2009) S289–S298.

[9] E. Long, Patterns of Participation in Canadian Literacy and Upgrading Programs: Results of a National Follow-Up Study., ERIC, 2001.

[10] Gartner, Gartner says worldwide pc, tablet and mobile phone combined shipments to reach 2.4 billion units in 2013, 2013. Http://www.gartner.com/newsroom/id/2408515 (accessed April 5, 2013).

[11] R. K. Megalingam, A. P. Rajendran, A. T. Solamon, D. Dileep, Edupada tablet based educational system for improving adult literacy in rural india, in: Technology Enhanced Education

(ICTEE), 2012 IEEE International Conference on, IEEE, pp. 1–5.

[12] K. Toukonn, The dynamic electronic textbook: Enhancing the student's learning experience, Kent State University (2011). MFA Thesis.

[13] S. P. Rushton, J. Eitelgeorge, R. Zickafoose, Connecting brian cambourne's conditions of learning theory to brain/mind principles: Implications for early childhood educators, Early Childhood Education Journal 31 (2003) 11–21. 10.1023/A:1025128600850.

[14] B. X. Chen, Hmh fuse pilot program, 2012. Http://www.hmheducation.com/fuse/pilot-1.php.

[15] S. Henderson, J. Yeow, ipad in education: A case study of ipad adoption and use in a primary school, in: System Science (HICSS), 2012 45th Hawaii International Conference on, pp. 78–87.

[16] A. L. Culén, A. Gasparini, ipad: a new classroom technology? a report from two pilot studies, INFuture Proceedings (2011) 199–208.

[17] J. W. Gentry, What is experiential learning, Guide to business gaming and experiential learning (1990) 9–20.

[18] S. Deterding, Situated motivational affordances of game elements: A conceptual model., 2011. Presented at Gamification: Using Game Design Elements in Non-Gaming Contexts, a workshop at CHI 2011. Retrieved from http://gamification-research.org/wp-content/uploads/2011/04/09-Deterding.pdf.

[19] Zamzee, New research shows zamzee increases physical activity by almost 60%, 2012. Http://blog.zamzee.com/2012/09/26/newresearchshowszamzeeincreasesphysicalactivitybyalmost60/ (accessed April 5, 2013).

[20] F. Barry, How online fundraising, gamification and social media helped raise over $2 million dollars in one day, 2012. Http://www.npengage.com/onlinefundraising/howonlinefundraisinggamificationandsocialmediahelpedraiseover2-milliondollarsoneday/ (accessed April 5, 2013).

[21] Gartner, Gartner says by 2014, 80 percent of current gamified applications will fail to meet business objectives primarily due to poor design, 2012. Http://www.gartner.com/newsroom/id/2251015 (accessed April 5, 2013).

[22] K. Wattanatchariya, S. Chuchuaikam, N. Dejdumrong, An educational game for learning wind and gravity theory on ios: Drop donuts, in: Computer Science and Software Engineering (JCSSE), 2011 Eighth International Joint Conference on, pp. 387 –392.

[23] F. Yan, A sunny day: Ann and rons world an ipad application for children with autism, in: M. Ma, M. Fradinho Oliveira, J. Madeiras Pereira (Eds.), Serious Games Development and Applications, volume 6944 of *Lecture Notes in Computer Science*, Springer Berlin / Heidelberg, 2011, pp. 129–138.

[24] S. Deterding, D. Dixon, R. Khaled, L. Nacke, From game design elements to gamefulness: defining "gamification", in: Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments, MindTrek '11, ACM, New York, NY, USA, 2011, pp. 9–15.

[25] M. Csikszentmihalyi, Finding flow: The psychology of engagement with everyday life, Basic Books, 1997.

[26] A. Dorling, F. McCaffery, The gamification of spice, in: Software Process Improvement and Capability Determination, Springer, 2012, pp. 295–301.

[27] J. Nakamura, M. Csikszentmihalyi, Flow theory and research, Handbook of positive psychology (2009) 195–206.

[28] J. Hamari, V. Eranti, Framework for designing and evaluating game achievements, Proc. DiGRA 2011: Think Design Play 115 (2011) 122–134.

[29] J. Keller, H. Bless, F. Blomann, D. Kleinböhl, Physiological aspects of flow experiences: Skills-demand-compatibility effects on heart rate variability and salivary cortisol, Journal of Experimental Social Psychology 47 (2011) 849–852.

[30] D. Johnson, J. Wiles, Effective affective user interface design in games, Ergonomics 46 (2003) 1332–1345.

[31] J. SimíEs, R. D. Redondo, A. F. Vilas, A social gamification framework for a k-6 learning platform, Comput. Hum. Behav. 29 (2013) 345–353.

[32] J. Klein, Y. Moon, R. W. Picard, This computer responds to user frustration, in: CHI'99 extended abstracts on Human factors in computing systems, ACM, pp. 242–243.

[33] K. D. Squire, Video games in education, Int. J. Intell. Games & Simulation 2 (2003) 49–62.

[34] R. N. Landers, R. C. Callan, Casual social games as serious games: The psychology of gamification in undergraduate education and employee training, in: M. Ma, A. Oikonomou, L. C. Jain (Eds.), Serious Games and Edutainment Applications, Springer London, 2011, pp. 399–423. 10.1007/978144712161-9_20.