Identifying Cognitive Load in a Computer Game:
An exploratory study of young children

Allan Fowler
College of Computing and Software Eng.
Kennesaw State University
Marietta, USA
afowle56@kennesaw.edu

Keith Nesbitt
School of Design Communication and IT
University of Newcastle
Newcastle, Australia
keith.nesbitt@newcastle.edu.au

Alessandro Canossa
School of Design
Royal Danish Academy of Fine Arts
Copenhagen, Denmark
acan@kadk.dk

Abstract—There has been considerable interest in measuring cognitive load. While researchers use retrospective think aloud methods to gauge an individuals cognitive load, others have expressed concerns about the reliability of think aloud methods. Moreover, others have expressed concerns about the reliability of using think-aloud methods with young children. This article provides evidence of non-invasive psycho-physiological tools that can be used to measure the cognitive load of young children. The authors employ eye-tracking technology to measure the players cognitive load through changes in endogenous blinks and fixations. Our process finds significant changes in cognitive load between childrens first and second attempts at playing puzzle-based computer games. The results of the study provide game developers with tools that indicate when users struggle with the information or content in a computer game.

Index Terms—Player cognition, endogenous blinks, fixations, eye tracking, cognitive load, games

I. INTRODUCTION

Computer or video games are used both for entertainment purposes and for education. Technologies such as computer games provide effective scaffolding, increase the opportunity for a learner to receive feedback, and to build both local and global communities of learners [40]. Although the debate continues about the value of computer games, the recent introduction of affordable non-invasive tools that provide an opportunity to measure cognition of the user while playing games [38]. These tools can provide researchers and game developers with qualitative and quantitative data on the user experience. This data has been used to improve the quality of the game and measure cognitive processes that have traditionally been difficult to measure accurately.

One theory that is used to measure human cognition is the Cognitive Load Theory [33], [44]. Cognitive Load Theory (CLT) is based on the understanding that cognitive load measures the amount of mental effort being used in the working memory of humans. Information that is being learned has an intrinsic cognitive load and therefore the amount of cognitive load can be used to determine how difficult it is to learn or perform [44].

Game usability researchers use a number of tools and techniques to understand the player experience for the purpose of improving the overall game play experience [34]. One of the tools used to help understand the game play experience is the Think Aloud Methods (TAM). TAM elicit concurrent or retrospective verbal feedback on the users experience. While TAM provide an efficient method of understanding the player experience which helps in understanding the core issues with the game, TAM have several limitations [12], [35], [48]. As certain cognition is unconscious, participants will generally not be capable of verbalizing some or all of their thought processes [12], [35]. Moreover, most cognition is faster than verbal processes and therefore participants could be thinking about a lot more than what they can verbalize [12]. Finally, as the user needs to verbalize their thoughts as they play the game, concurrent thinking aloud can interrupt the game being played and this interruption could result in the participant being distracted while playing the game. The distraction could also add more cognitive load to the participant.

Using non-invasive psycho-physiological tools that can monitor cognitive load allows researchers to provide qualitative data on the play experience. The recent introduction of affordable high-resolution eye-trackers that can remotely track the cumulative fixation points, saccades (a rapid movement of the eye between fixation points), and blinks can provide detailed information about usability issues [23]. These remote systems can be placed in front of the participant and therefore, can monitor the user experience without interruption.

II. THEORETICAL BACKGROUND

A. Physiological indicators of cognition

What part of the screen do users look at, how long they look, and how many times they look at that particular object, can provide an indicator of what the user is thinking about [46]. Furthermore, this can also indicate where the user was having difficulty with that particular part of the game. Through identifying the cumulative fixation time (or gaze points), it is possible to determine the areas of attention [36]. It has been suggested that there is a relationship between what the eyes fixate on and cognition [27]. In order to comprehend visual information, the eyes fixate on areas that are surprising, significant, important, or need further investigation [25], [27]. Eye blinks also provide an indicator of human cognition and are possible indicators of cognitive load [17], [41]. Eye blinks are either a response to a startle, a voluntary action or due
to some other internal cause (an endogenous action) [41]. A startle blink is typically the result of when an object is about to enter the eye, although the reflex can also be triggered by bright flashes, loud noises or unexpected tactile inputs [3]. The amplitude of this type of blink has been used to gauge the affective state of players in computer games [2], [29], [31]. A voluntary blink is a direct result of the conscious decision to blink. An endogenous blink is caused by an unconscious process that is a result of perception, reaction or information processing [32] and is, therefore, indicative of player cognition (Figure 1 [45]). Eye blinks are frequently measured with a blink electrooculogram that requires the placement of electrodes on the extra-ocular muscles of the participants [3]. While this tool can produce reliable data on the latency, amplitude, and duration of each blink, it is intrusive to the participant. As the duration of the endogenous blink is unique [32], it is possible to identify this behavior using a high-frequency eye tracker [18].

The minds-eye theory states that where the eyes are focused on is what the person is thinking about [5], [26], [28]. The eyes tend to fixate when an object (or event) is surprising, significant, important, or needs further investigation [26], [30]. The long fixations are objects that the person is investigating or trying to understand (or problem solving). Longer fixations are considered to be evidence of increased cognition [15], [22], [32].

B. Measuring cognition in video games

Computer games have the potential to alter behavior, produce learning, and promote brain plasticity [1], [21]. Playing action computer games has been linked to enhancements in basic perceptual tasks including those involving peripheral vision [6], visual acuity [20], and temporal processing [11]. Moreover, playing action computer games appears to improve visual attentional skills [16]. Furthermore playing action computer games can improve higher order cognitive functions like multitasking [8], [42] and has the potential to improve or enhance working memory [43]. Jie and Clarke [24] used a 250Hz eye tracker to understand the reaction times (visual pursuit) in a custom made action video-game. The authors reported that participants (N=5) that played the difficult game demonstrated longer reaction times.

Nasr and Yan [13] used a 500Hz head-mounted eye tracker to collect the participants visual attention patterns in two commercial action games that required different reaction speeds and used a different camera model (first-person versus a third-person perspective). The authors found that the participants (N=6) demonstrated significant variations in visual patterns in different interventions. Shute, Ventura and Ke [38] found that a problem solving game may also improve or enhance certain aspects of cognitive functions. For example, through playing Portal 2 [9], players improved in problem solving and spatial reasoning skills [38].

C. Hypotheses

Based on the literature [17], [26], [28], [41] the authors predict that (H1) endogenous blinks and long eye fixations are indicators of cognitive load.

III. METHODOLOGY

The studies were conducted in a laboratory environment to minimize external distractions. The protocol was explained to the participants (and their parents) and informed consent (and assent), and initial demographic data was obtained (IRB number -redacted-). The computer monitor was positioned directly in front of the participant, and the SMI RED 500 eye tracker was placed under the screen (see figure 2). Participants were selected to play either World of Goo [4] or Bad Piggies [14].

A. World of Goo

The video game World of Goo [4] was chosen as the treatment because it was not overly popular and, therefore, potentially few participants will have had prior experience in playing it. Furthermore, the game is non-violent (which was considered important when asking young children to play it), and it does not require advanced reaction speeds. According
to Shute and Kim [37], World of Goo incorporates the physics concept of static equilibrium. The other implicit physics concept was the importance of building sound structures [10]. Through a pilot study, the researchers also identified that the additional physics concept embedded in the game was the concept of force (gravity, wind, or buoyancy). Furthermore, success in the game also depended on learning the basic principles of construction (the importance of strong foundations, the importance of support structures, and/or the importance of level structures). Furthermore, there is an opportunity to improve analytic thinking skills [10], [37]. World of Goo is a physics-based puzzle/construction game [10]. The basic premise of World of Goo (depending on the level) is to construct a tower, bridge, or chain to enable the goo (or gooballs) (the protagonists) to get close enough to the pipe. When the structure gets close to the pipe, the gooballs are extracted into an extraction pipe to join the liquefied goo. However, the forces of gravity and wind, spikes, bog (swamp), fire, cogs, or machines challenge the player to build something that is strong enough to resist the antagonist(s) but is also tall or long enough to reach the extraction pipe. The authors concluded that this game would require considerable cognitive load for the participants to complete.

B. Bad Piggies

The video game, Bad Piggies [14] was chosen for the control group because although it is a problem solving game unlike World of Goo it does not involve constructing towers. The basic premise of Bad Piggies is to build a vehicle (a wooden car or airplane) to transport a pig (the protagonist) across a variety of terrain (an antagonist) to the end of a predefined path to a nest. In the more advanced levels, there are also characters from the game Angry Birds that will try (by throwing something) to stop the pig getting to the egg. If the user gets the pig to the nest egg, the end goal is achieved. However, additional rewards are available for: breaking/not breaking the vehicle, bringing the King Pig in the vehicle, not using a specific vehicle part, collecting star crates, getting the pig to the destination on time, carrying an egg in the vehicle, and/or carrying cakes or treats. There are 168 story levels within the game. Movement to the higher levels is dependent on obtaining two or three stars in each of the lower levels. If three stars are obtained, the user may also be re To ensure the frequency of endogenous blinks in the study, the researchers used the SMI BeGaze(TM) eye-tracking software to identify the blinks and fixations of each participant. Each blink was categorized based on the speed into three frequency ranges. The endogenous blink was operationalized as being between 100 and 300 milliseconds (ms), which is within the range identified in the literature [15], [23], [50]. The eye tracker used (SMI RED500) is a bi-ocular device, and therefore, it identifies the closing of the eyelid in both or either eyes. The study focused on blinking (both eyelids closed) and not winking (one eyelid closed). Therefore, comparisons can be made between the two exposures.

IV. RESULTS

Upon completion of the experiments, the researchers used the SMI BeGaze(TM) eye-tracking software to identify the blinks and fixations of each participant. Each blink was categorized based on the speed into three frequency ranges. The endogenous blink was operationalized as being between 100 and 300 milliseconds (ms), which is within the range identified in the literature [15], [23], [50]. The eye tracker used (SMI RED500) is a bi-ocular device, and therefore, it identifies the closing of the eyelid in both or either eyes. The study focused on blinking (both eyelids closed) and not winking (one eyelid closed). Therefore, if the software classified the blink with both eyes, the data was included. Due to the variation of the duration of game play, the researchers normalized the data to provide a consistent standard for comparison.

The percentage of endogenous blinks was calculated by dividing their number of endogenous blinks by the total number of blinks (Table 1). The percentage of fixations over 600ms was calculated by classifying all fixations by the duration. Overall the participants showed a mean reduction of 22% in the percentage of endogenous blinks from the first attempt ($M = 0.54, SD = 0.13$) to the second attempt ($M = 0.32, SD = 0.16$) of the game. Only one participant (K-10) demonstrated an increase in the percentage of endogenous blinks in the second attempt at the game. A one-way between subjects ANOVA was conducted to compare the percentage blinks in the childrens
first and second attempt at the game. There was a significant difference in the percentage blinks between attempts at the p<0.05 level for the two conditions [F(1, 10) = 11.41, p = 0.003]. The data had a skewness of 1.26 and kurtosis of 1.84.

Participants also showed a mean reduction of 4% in the percentage of fixations greater than 600ms that occurred in participants between the first attempt (M=6.93, SD=5.09) and second attempt (M=2.87, SD=2.82) of the game. Only one participant (K-3) demonstrated an increase in the percentage of endogenous blinks in the second attempt at the game. A one-way between subjects ANOVA was conducted to compare the percentage of fixations over 600ms in the childrens first and the second attempt at the game. There was a significant difference between attempts at the p<0.05 level for the two conditions [F(1, 10) = 4.90, p = 0.040]. The data had a skewness of 1.59 and kurtosis of 2.58.

The authors next examined any changes in performance for participants between their first and second attempt. The in-game performance of players of Bad Piggies is indicated by the number of stars (0-3), and participant performance showed no indication of changing between attempts. In-game performance in World of Goo is measured by a broader and finer resolution range of measures, namely the number of Gooballs collected, the number of moves required to collect them and the time taken to complete the level. Overall all the participants that played World of Goo improved in finishing the level on most of these measures. Although participant K-01 did not complete (DNC) the level in both attempts, this participant still improved in the time it took to run completely out of moves.

V. DISCUSSION

The two measures of cognitive load, namely the percentage of endogenous blinks and the percentage of long fixations were obtained using non-invasive eye tracking technology. As noted, the authors identified a significant reduction in endogenous blinks in the children. This indicates that there was a reduction in the cognitive load in the second attempt. Moreover, there was a notable difference between the percentage of endogenous blinks and long eye fixations in the two games. This result validates the work of Fogerty & Stern [17], Stern, Walrath, & Goldstein [41] and Orchard & Stern [32] in that long eye fixations and endogenous blinks provide valuable insight into human cognition.

The results support the first hypothesis (H1) that cognitive load appears to be measurable through physiological measures. This is evident from the higher percentage of both fixations and endogenous blinks in the World of Goo, which is a more complex game as it required more advanced problem solving.

It needs to be noted that several causes can trigger the reduction of cognitive load, having learned the task is definitely one, but also the presentation of a familiar set of stimuli can account for decreased cognitive activity for at least two reasons: a familiar stimulus does not present elements of novelty, furthermore being the challenge also inferior there is a risk of boredom onset. The authors also identified significant reductions in long fixation. The children reduced the quantity and frequency of endogenous blinks and the number of long fixations between each exposure. The children appeared to be learning to play the game (both perceptual [19] and conceptual learning [49]).

Another aspect to take into consideration when interpreting the results is the wide variety of motivations and problem solving strategies that are usually ascribed to individual differences. Established knowledge regarding variability in cognition looks at inter-individual (between different people) and intra-individual (within a given person) variability [39]. The variability in cognition shown both within each of the two groups and across the two groups somewhat mirrors previous experiments. It is, therefore, easy to explain unexpected outliers like K-10. Recent research has demonstrated how personality and individual differences can account for variation and diversity in behavior more than demographics [7], [47].

VI. CONCLUSION

In the experiment, the authors collected two measures of cognitive behavior, namely the percentage of endogenous blinks and the number of eye fixations greater than 600ms. These measures were obtained using non-invasive eye tracking technology. The participants were randomly allocated to either play Bad Piggies or World of Goo. None of the participants were familiar with these games. All participants played their allocated game twice so the authors could examine changes in their cognitive load between the two attempts. The authors then examined changes in the two cognitive load, endogenous

<table>
<thead>
<tr>
<th>Game Played</th>
<th>Participant</th>
<th>% of Endogenous Blinks</th>
<th>% of Fixations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Attempt 1</td>
<td>Attempt 2</td>
</tr>
<tr>
<td>World of Goo</td>
<td>K-01</td>
<td>66.67</td>
<td>25.61</td>
</tr>
<tr>
<td></td>
<td>K-02</td>
<td>60.68</td>
<td>22.72</td>
</tr>
<tr>
<td></td>
<td>K-03</td>
<td>25.00</td>
<td>11.11</td>
</tr>
<tr>
<td></td>
<td>K-08</td>
<td>62.53</td>
<td>47.77</td>
</tr>
<tr>
<td></td>
<td>K-10</td>
<td>64.65</td>
<td>66.67</td>
</tr>
<tr>
<td></td>
<td>K-11</td>
<td>58.46</td>
<td>39.46</td>
</tr>
<tr>
<td>Bad Piggies</td>
<td>K-04</td>
<td>43.24</td>
<td>24.00</td>
</tr>
<tr>
<td></td>
<td>K-06</td>
<td>40.91</td>
<td>28.57</td>
</tr>
<tr>
<td></td>
<td>K-07</td>
<td>56.25</td>
<td>30.00</td>
</tr>
<tr>
<td></td>
<td>K-09</td>
<td>60.68</td>
<td>25.67</td>
</tr>
</tbody>
</table>

TABLE I

EYE BLINK AND FIXATION RESULTS
blinks and fixations greater than 600ms between their first and second attempt. In the study both of these measures were significantly reduced between the first and second attempt. The authors also looked at changes in player performance between their first and second attempt. The in-game performance measure (3 stars) used in Bad Piggies did not provide a significant resolution to indicate any change in performance. In the case of World of Goo, we were able to obtain three performance measures (number of gooballs collected, the number of moves and time to complete). Analyzing these three measures provided an indication that the performance of both the children (n=6) World of Goo improved their performance between attempts. The non-invasive eye-tracking approach can measure changes in cognitive load while the children are playing a computer game. The significant changes found in the two measures of cognitive load and also the improved performance of participants is suggestive of learning that has occurred between attempts.

The sample size may raise questions about the generalizability of these findings. Similar studies have used a smaller sample size. [13] used six participants and [24] used five participants. Drachen, Canossa, and Srensen [51] suggest that an initial study of three participants is a reasonable sample size for an initial (pilot) usability study.

From the evidence the authors conclude that the methods can be used for evaluating where the user is in a flow state (or not). Moreover, these methods can be used for in user testing experiments and or, testing if the game is playable (or not). Moreover, these methods can be used for evaluating where the user is in a flow state. The in-game performance measures (number of gooballs collected, the number of moves and time to complete). Analyzing these three measures provided an indication that the performance of both the children (n=6) World of Goo improved their performance between attempts. The non-invasive eye-tracking approach can measure changes in cognitive load while the children are playing a computer game. The significant changes found in the two measures of cognitive load and also the improved performance of participants is suggestive of learning that has occurred between attempts.

ACKNOWLEDGMENT

This research was conducted under the auspices of the Northeastern University Ethics and Internal Review Board (IRB Number: 13/10/13). This research was partially funded by the Auckland University of Technology Research Grant. The authors would like to acknowledge the assistance of SMI, who provided the RED500. We thank 2D Boy, the developers of World of Goo for their support and permission to use images from their game. We also would like to acknowledge the parents and children who took part in this study.

REFERENCES


