Modeling Visual Attention of Students Playing an Educational Game for Physics

Celesamae T. Vicente[a], Ma. Mercedes T. Rodrigo[a], Hong-Fa Ho[b], Guan-An Chen[b]

[a]Department of Information Systems and Computer Science, Ateneo de Manila University
[b]Department of Electrical Engineering, National Taiwan Normal University

celesamae2010@gmail.com, mrodrigo@ateneo.edu,
jackoho@ntnu.edu.tw, crackereidolon@gmail.com

Abstract. This study models the visual attention of students as they think of a solution to a problem within an educational game for Physics. Participants were given time to view a hint followed by a static image of the game problem. Upon viewing the problem, students were instructed to think of a solution using the hint while an eye tracker recorded eye movement data. After viewing the problem, participants played the actual game level. The game awarded gold, silver, or no badges to participants depending upon their performance. When analyzing the relationship between the eye movement and performance, the findings were that participants who earned gold, silver, and no badges had different orders of fixating on the regions of interest while thinking of a solution to the problem and participants who had better performance fixated earlier on the regions where the solutions should have been drawn.

Keywords: Attention, Physics Playground, Eye Tracking

1 Introduction

Attention refers to concentration given to one phenomenon, to the exclusion of others [10]. It requires tuning out of other stimuli so that a person can apply himself or herself to the phenomenon of interest [8]. The focus of this paper is attention in the context of learning. Research has shown that student performance levels are associated with endurance or attention span [4]. Fluency or accuracy of learner performance in significant durations has been linked with higher or sustained attention [4]. In this study, we examine the relationship between attention and performance among students using Physics Playground (PP), an educational game for designed to help secondary students understand Newtonian Physics [14]. PP helps students better understand balance, mass, conservation and transfer of momentum, gravity and potential kinetic energy [13]. PP has 80 levels. The goal in each level is for students to guide a green ball to a red balloon by drawing machines such as the inclined plane or ramps, levers, pendulums, and springboards. PP awards badges to the players depending upon the number of objects the player used. A gold badge is given when players solve a problem at or below a par number of objects for the problem. Silver badges, on the other hand, are awarded when the player solves the problem with a greater than par number of objects. All player actions are recorded in the log files [13].

PP's impact on student achievement is varied. Studies in the US [3] showed that PP is able to augment students' understanding of qualitative physics. In the Philippines, however, PP did not result in any learning gains [1][2]. Why is it that so? This study investigates the role of attention (or lack thereof) among Philippine students playing PP using eye movement.

Just and Carpenter provide some of the earliest studies on attention and eye-movements. They coined the "eye-mind" hypothesis which pertains to the notion that eye-movement is a "on top of the stack" indicator of cognitive process and that it can give a trace of where attention is directed [9]. Because of this, researchers have developed means of quantifying attention. One of which is the chosen method in conducting this study, eye-tracking. Eye-tracking is a method that quantifies eye movements relative to the current line of sight of a participant in a stimulus at a given time. When eye movements are quantified, they are broken down into two major metrics, saccades and fixations. Saccades are sudden changes of eye gazes between points of fixations [11]. While fixation is maintaining of eye gaze at a certain position indicating a person's intention of what to interact with [16], what current task a person is working on [15] and is a proxy indicator of attention[5]. Metrics on fixation and gaze durations have been used in attempt to find answers to research questions of this study. This paper attempts to determine the relationship between eye gaze and student performance.

2 Methodology

2.1 Participants

A total of 30 high school students were recruited from different parts of the Philippines for this experiment. There were 19 male and 11 female participants, 10 of whom were from Luzon, 11 from Visayas, and 9 are from Mindanao. The mean average of participants was 15.5 years and participants were in grade levels 9 to 11. Participants typically spent 2.9 hours a day playing video games and 3.6 hours a day watching television. Participants' Physics grades' were 87.9 out of 100 on average. Participants were asked to play the tutorial portion of PP to become acquainted with the game mechanics. After the tutorial, participants viewed static images of a hint and the problem to think of a solution while the eye-tracker was recording their eye movements. Details on this will be discussed in the next section. Participants then played the PP level they viewed. Participants were given a maximum of 3 minutes to solve each level. Finally, participants took a post-test that was isomorphic to the pre-test. Comparison of pre-test and post-test scores showed no improvement after using the PP.

2.2 Material

There were two types of stimuli used in this experiment. The first type (Figure 1, left) contained a hint regarding the simple machine that was most relevant to the solution.

The second type (Figure 1, right) was the static image of the pre-selected game level problem from PP named Scale. Participants were instructed to look at the PP problem while they were thinking of how to bring the green ball to the red balloon. While they were doing this, the eye-tracker recorded their eye-movements. The regions of interest (ROI) of each stimuli were later defined according to the types described below:

1. Hint - Simple machine hint that was most relevant to the solution
2. Instruction - Instruction about what keys to press to go to the next picture
3. Starting Point - Initial position of the green ball
4. Target - Location of the red balloon
5. Solution Space - Where the simple machine solution is ideally drawn
6. Decision Factor - Objects in the stimulus that implicitly contribute to the decisions participants make in solving the problem
7. Travel Path - The green ball passes through this area to reach the red balloon
8. Existing Objects - Objects that are not explicitly essential in solving the problem

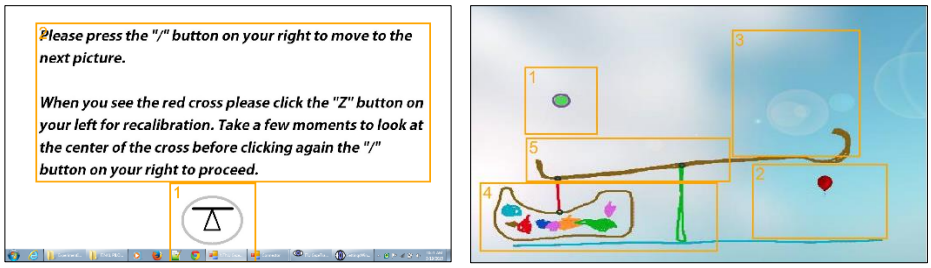

Fig. 1. Stimulus 1 on the left is the Lever Hint for Stimulus 2 PP Scale level on the right.

There were 2 types of ROIs in the hint, and 6 types of ROIs in the PP level. On the Lever hint, the ROIs defined are: 1) R1 – Hint, 2) R2 - Instructions. The static image for PP game level Scale has the following ROIs: 1) R1 – Starting Point, 2) R2 – Target/Solution Space, 3) R3 – Solution Space, 4) R4 –Decision Factor and 5) R5 – Travel Path. The regions that were not defined among these categories fell under the "others" classification. These defined ROIs are the essential points of investigation and comparison during the analysis. The ROI definitions on the PP game level were assigned on the basis of where solutions were drawn. Figure 2 shows two ways of solving the problem. Objects can be drawn on R3 or R2 to offset the weight of the Scale and bring the green ball to red balloon.

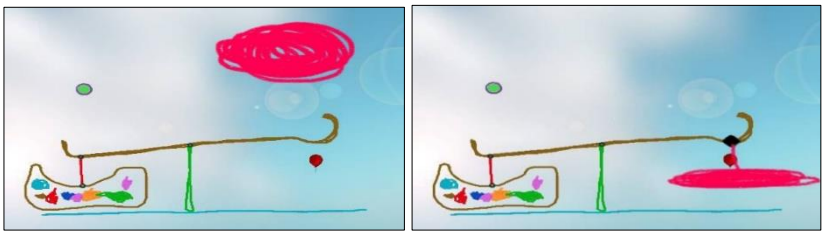

Fig. 2. Lever solutions to bring the green ball and to the red balloon.

2.3 Instrument

The eye movement data were recorded using the EyeNTNU-120 eye tracker that has been used to study visual attention on reading process on integrated circuits in [6] and on how women perceive handbags in [7] . Figure 3 shows the set-up during the data gathering. Participants were asked to place their chins on the chin-rest while the eye-camera was directed at one of the participant's eyes. As the participant viewed the onscreen stimuli, the eye-tracker recorded and mapped participants' eye movements in regions. This device had a sampling rate of 120 Hz and an error rate of less than 0.3 degrees given that the participants were less than 60 centimeters away from the computer screen. The four vital metric variables provided by the system and have been the basis for data analysis are the following:

- Total Contact Time (TCT) – total time in milliseconds a participant gazed on ROI.
- Number of Fixations (NOF) –the number of times the participant fixated on ROI
- Duration of First Fixation (DFF) – the total time in milliseconds that the first fixation on ROI lasted, and
- Latency of First Fixation (LFF) – the time when the first fixation on ROI occurred.

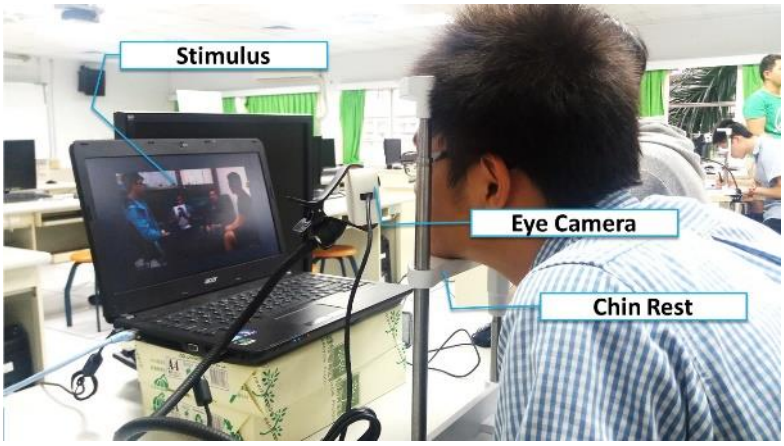

Fig. 3. EyeNTNU-120 Camera and Chin-rest Set-up

User eye gaze data and user action logs from PP were later synchronized in order to investigate the relationship between eye gaze data with user performance based on badges earned.

3 Results and Discussions

Out of 30 participants, 20 solved the problem. Six participants earned gold badges and the other 14 received silver badges. Ten participants did not earn any badges. The succeeding sections discuss the analysis of the metric values of the ROIs for both

Stimulus 1 and Stimulus 2 and the consolidated ROI grouped by badges earned. Percentages were used to determine the rate of TCT and NOF to normalize data.

3.1 Total Contact Time (TCT)

Stimulus 1 – Lever Hint for PP Game Level Scale.
As seen in Table 1, all participants spent more time gazing at R2- Instructions than in R1- Hint. For R1, gold badge earners had the highest percentage of TCT (22.71%) followed by those who did not earn any badge (12.6%) and the silver badge earners (7.1%). There is a statistically significant difference on the TCT of R1 among the groups based on one-way ANOVA ($F(2,27)=6.7$, $p = 0.004$). A Tukey post-hoc test showed that TCT of gold badge earners were significantly higher than those who earned silver badge ($7.1\%\pm6.9\%, p = 0.003$) and marginally higher than those who did not earn any badge ($12.6\%\pm8.4\%, p=0.083$). For R2 the instructions, silver badge earners had the highest TCT (87.9%), followed by those who did not earn any badge (82.8%) and then the gold badge earners (72.8%). There is a statistically significant difference on the TCT of R2 among the groups based on one-way ANOVA $F(2,27)=5.6, p=0.010$. Gold badge earners had significantly lower TCT on the instructions than those who earned silver badge ($8.8\%\pm7.1\%$, $p=0.007$).

BADGE	Hint	Instructions	ORDER
	R1	R2	From Greatest
GOLD	2 (22.7%)	1 (72.8%)	R2 > R1
SILVER	2 (7.1%)	1 (87.9%)	R2 > R1
NO BADGE	2 (12.6%)	1 (82.8%)	R2 > R1

Table 1. TCT percentage per ROI on Stimulus 1 – Hint for PP Game Level Scale

Stimulus 2 – PP Game Level Scale.
The order of ROI from the greatest to the least TCT grouped by badge earned in Table 2 shows that Gold badge earners spent the highest time in R2, R4, and R3 which have ROI types Target/Solution Space, Decision Factor and Solution space. Silver badge earners and those who did not earn any badge on the other hand spent less time looking at the top three ROI types than gold badge earners. Results also show that the Gold badge earners had the highest TCT in R2 (solution space) (28.51%) followed by the silver badge earners (16.02%) and those who did not earn any badge (8.82%).

BADGE	Starting Point	Target /Solution Space	Solution Space	Decision Factor	Travel Path	ORDER
	R1	R2	R3	R4	R5	From Greatest
GOLD	5 (11.2%)	1 (28.5%)	3 (14.3%)	2 (19.5%)	4 (13.4%)	R2 > R4 > R3 > R5 > R1
SILVER	1	3	5	4	2	R1> R5 > R2 >

Table 2. Percentage and Order of TCT from the Greatest to the Least Grouped by Badge

	(18.7%)	(16.0%)	(12.7%)	(15.0%)	(17.7%)	R4 > R3
NO BADGE	2 (20.0%)	4 (9.0%)	5 (5.2%)	3 (17.0%)	1 (26.7%)	R5 > R1 > R4 > R2 > R3

Consolidated ROI.

R1 of Stimulus 1 and R1 to R5 of Stimulus 2 were consolidated into a single ROI. The TCT of gold, silver, and no badge participants were averaged and compared. Gold, no badge, and silver badge earners spent 18.3%, 15%, and 14.5% of their TCT looking at this consolidated region respectively. The gold badge earners were more engaged in the regions identified to be crucial in accessing information in solving a problem. There was a statistically significant difference between groups as determined by one-way ANOVA ($F(2,27)=4.307$, $p= 0.024$). A Tukey post-hoc test revealed that the consolidated TCT of gold badge earners was statistically higher than those who earned silver badge ($14.5\%\pm2.6\%$,$p=0.021$) and marginally higher than those who did not earn any badge ($15.0\%\pm3.0\%$, $p=0.066$). There was no statistically significant difference between the gold badge earners and silver badge earners ($p=0.893$).

3.2 Number of Fixations (NOF)

Stimulus 1 – Lever Hint for PP Game Level Scale.
Table 3 shows the percentages of NOF per ROI in Stimulus 1. Gold badge earners had the highest NOF percentage on R1 at 20% followed by those who did not earn any badge at 13.1% and finally by silver badge earners at 7.2%. There is a statistically significant difference between groups determined using by one-way ANOVA ($F(2,27)=5.367$, $p=0.011$). Gold badge earners had significantly more NOF on R1 than silver badge earners ($7.2\%\pm6.3\%$,$p=0.009$).

BADGE	Hint	Instructions	ORDER
	R1	R2	From Greatest
GOLD	2 (20.0%)	1 (76.0%)	R2 > R1
SILVER	2 (7.2%)	1 (88.1%)	R2 > R1
NO BADGE	2 (13.1%)	1 (83.5%)	R2 > R1

Table 3. NOF Percentage on Stimulus 1 – Hint for PP Game Level Scale

Stimulus 2 – PP Game Level Scale.
Table 4 shows the NOF percentage per ROI in stimulus 2. Ranking NOF percentages from the greatest to the least show that silver badge earners and those who did not earn any badge had the same order. Both groups had the highest fixation on ROIs R5, then R1, then R4. ROIs types R2 and R3 spaces had the least NOF percentage. While gold badge earners on the other hand, had the highest NOF percentage on ROIs R2, R4, and R3 and with the least NOF on the R5 and R1. This is indicative that gold badge earners had higher attention on ROI types Target, Solution Space, and Decision

Factors. While both silver badge and no badge earners had the least attention on these ROI types. There is a marginal significant difference between groups in R2-target location of red balloon, as determined by one-way ANOVA($F(2,27)=2.782, p=0.08$). A post-hoc Tukey test showed that gold badge earners have marginally higher NOF than those who did not earn any badge (8.8%±8.8%,$p=0.066$).

Badge	Starting Point	Target /Solution Space	Solution Space	Decision Factor	Travel Path	Order (from greatest to least NOF)
	R1	R2	R3	R4	R5	
GOLD	5 (11.3%)	1 (28.6%)	3 (14.3%)	2 (16.3%)	4 (14.1%)	R2 > R4 > R3 > R5 > R1
SILVER	2 (18.1%)	4 (14.2%)	5 (12.7%)	3 (15.0%)	1 (18.8%)	R5 > R1 > R4 > R2 > R3
NO BADGE	2 (19.4%)	4 (8.8%)	5 (5.2%)	3 (17.5%)	1 (26.3%)	R5 > R1 > R4 > R2 > R3

Table 4. NOF percentage per ROI on Stimulus 2 – PP Game Level Scale.

Consolidated ROI.

R1 of Stimulus 1 and R1 to R5 of Stimulus 2 were consolidated into a single ROI. The NOF of gold, silver, and no badge participants were averaged and compared. Gold, no badge, and silver badge earners spent 17.4%, 15%, and 14.5% of their NOF looking at this super region respectively. There was a marginal significant difference between groups as determined by one-way ANOVA ($F(2,27)=2.896$, $p=0.073$). A Tukey post-hoc test revealed that the gold badge earners had marginally significant higher NOF than those who earned silver badge (14.5%±2.4%,$p=0.061$). There is no significant difference between the NOF of those who did not earn any badge and those who earned gold badge ($p=0.180$) and those who earned silver badge ($p=0.858$).

3.3 Latency of First Fixation (LFF) and Duration of First Fixation(DFF)

Figure 4 shows that all participants parsed the ROIs in Stimulus 1 in the same manner. That is, the first fixation was R2 (Instructions), followed by R1 (Hint). On average, Gold badge earners fixated on R1 at the 9.31 second mark, then followed by those who did not earn any badge at the 10.04 second mark and then by silver badge earners at the 16.51 second mark. The DFF orders on the other hand were the same for silver badge earners and those who did not earn any badge where in DFF on R2 was greater than R1. In addition, while the Gold badge earners had the earliest LFF on R1, they also fixated on this region the longest. This indicates that they fixated earlier on the hint and looked at it the longest compared to other participants. Table 5 shows the LFF (rank according to time access, from the earliest LFF) and DFF (rank according to length of time first fixation happened, from the greatest DFF) results for stimulus 1.

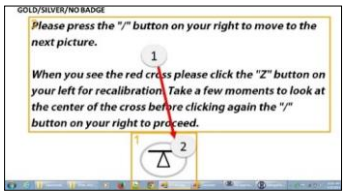

Fig. 4. LFF for Stimulus 1

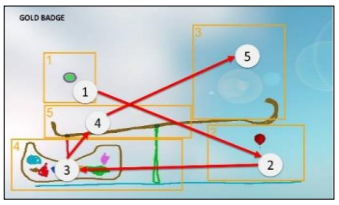

Fig. 5. LFF for Stimulus 2 – Gold Badge Earners

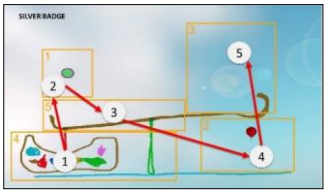

Fig. 6. LFF for Stimulus 2 – Silver Badge

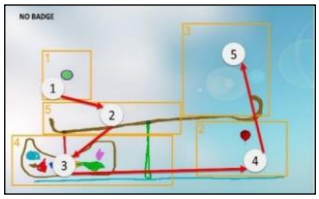

Fig. 7. LFF for Stimulus 2 – No Badge

BADGE	Metric	Hint	Instruction	ORDER
		R1	R2	
GOLD	LFF	2 (9.3s)	1 (0.3s)	R2 – R1
	DFF	1 (84.5ms)	2 (46.8ms)	R1 > R2
SILVER	LFF	2 (16.5s)	1 (0.3s)	R2 – R1
	DFF	2 (41.45ms)	1 (43.4 ms)	R2 > R1
NO BADGE	LFF	2 (10.0s)	1 (0.5s)	R2 – R1
	DFF	2 (38.4ms)	1 (39.9 ms)	R2 > R1

Table 5. LFF and DFF Ranking and Time Averages for Stimulus 1 – Lever Hint

Figures 5 to 7 show the LFF on Stimulus 2 of participants that earned Gold badges, Silver badges,and no badges. Table 6 shows that all groups of participants have different orders LFF and DFFs for Stimulus 2. In addition, Results of LFF showed that Gold badge earners accessed the ROIs that are solution spaces earlier, at 4.80 second mark, whereas those who did not earn any badge at accessed the same ROI at the 8.15 second mark. Interestingly, Gold badge earners accessed the solution space earlier but spent the least time on these spaces themselves. This implies that these participants arrived at the solutions earlier and faster than those who did not earn any badge.

Badge	Met-ric	Star-ting Point	Target /Solution Space	Solu-tion Space	Decision Factor	Travel Path	ORDER
		R1	R2	R3	R4	R5	
Gold	LFF	1 0.8s	2 3.1s	5 6.5s	3 5.4s	4 0.8s	R1 – R2 – R4 – R5 –R3
	DFF	2	4	5	1	3	R4 > R1 > R5 >

		52.5ms	40.0ms	26.3ms	71.2ms	44.0ms	R2 > R3
Silver	LFF	2 2.6s	4 4.7s	5 6.7s	1 2.5s	3 0.6s	R4 – R1 – R5 – R2- R3
	DFF	2 51.4ms	3 45.3ms	1 52.1ms	5 29.9ms	4 44.0ms	R3 > R1 > R2 > R5 > R4
No Badge	LFF	1 2.9s	4 6.7s	5 9.7s	3 4.5s	2 1.8s	R1 – R5 –R4 – R2 –R3
	DFF	5 39.1ms	2 48.8ms	4 40.4ms	3 48.1ms	2 52.8ms	R5 > R2 > R4 > R3 > R1

Table 6. LFF and DFF Ranking and Time Averages for Stimulus 2 – PP Game Level Scale

4 Conclusion and Future Work

Results show that students that had good performance paid more attention on the regions essential for solving the physics problem while they were still thinking of a solution. Time sequence in parsing the problem also showed that participants that performed better assessed the problem first by looking at the initial position of the green ball and then looking at the region of red balloon target and the solution space. Then they looked at objects that are contributory to solving the problem and then looked for an alternate solution. This is indicative that participants that had a good performance assessed what needs to be done to solve the problem and had a goal in mind on the first parsing through the regions the of interest. The participants that had a bad performance did not have the same sequence of parsing. In addition, participants who solved the problem accessed the solution spaces almost half the time earlier than those who did not earn any badge. This is indicative that students that earned badges had thought of solutions faster and earlier than those who did not solve the problem. Investigation on other stimuli in PP with different difficulty levels will be done in the future in order to know if a trend on attention levels can be discovered.

Acknowledgement

We thank the Ateneo de Manila University, the Ateneo Laboratory for the Learning Sciences, the National Taiwan Normal University (NTNU), the College of Technology and Engineering of NTNU, Dean Kuan-Chao Yu, Professor Chung-Yen Su, Juan Miguel Andres, Ryan S. Baker, Valerie Shute, Matthew Ventura, Matthew Small, Jaclyn Ocumpaugh, Jessica Sugay, Juliana Alexandra Andres, Michelle Banawan, Yancy Vance Paredes, and Nicko Regino Caluya.

We also thank the Department of Science and Technology's (DOST) Philippine Council for Industry, Energy, and Emerging Technology Research and Development for the grant entitled "Stealth assessment of student conscientiousness, cognitive-affective states, and learning using an educational game for Physics" and also the DOST-Engineering Research and Development Technology's Sandwich Program

Grant for Internship at National Taiwan Normal University that made this research project possible.

References

1. ANDRES, J. M. L AND RODRIGO, M. M. T. 2014, November. Learning and Affect Trajectories Within Newton's Playground. In 3rd International Workshop on ICT Trends in Emerging Economies, Nara, Japan
2. ANDRES, J. M. L AND RODRIGO, M. M. T. 2013 The Incidence and Persistence of Affective States While Playing Newton's Playground.
3. BAKER, R. S., D'MELLO, S. K., RODRIGO, M. M. T., AND RRAESSER, A. C. 2010. Better to be frustrated than bored: The incidence, persistence, and impact of learners' cognitive–affective states during interactions with three different computer-based learning environments. International Journal of Human-Computer Studies, 68(4), 223-241.
4. BINDER, C., HAUGHTON, E., & VAN EYK, D. 1990. Increasing endurance by building fluency: Precision teaching attention span. Teaching Exceptional Children, 22(3), 24-27.
5. BOLT, R. A. 1990, March. A gaze-responsive self-disclosing display. In Proceedings of the SIGCHI conference on Human factors in computing systems (pp. 3-10). ACM.
6. HO, H. F. (2013). Reading process of integrated circuit layout debugging: Evidence from eye movements. In *Advanced Materials Research.* (Vol. 787, pp. 855-860).(Advanced Materials Research; Vol. 787).10.4028/www.scientific.net/AMR.787.855
7. HO, H. F. (2014). The effects of controlling visual attention to handbags for women in online shops: Evidence from eye movements. *Computers in Human Behavior*, *30*, 146-152. 10.1016/j.chb.2013.08.006
8. JAMES, W. 1890. The Principles of Psychology, v. 1, New York: Henry Holt and Co. Reprinted in 1950.
9. JUST, M. A., & Carpenter, P. A. (1976). Eye fixations and cognitive processes. Cognitive psychology, 8(4), 441-480.
10. MATLIN, M. W. 2005. Cognition. Hoboken.
11. SALVUCCI, D. D., & GOLDBERG, J. H. 2000, November. Identifying fixations and saccades in eye-tracking protocols. In Proceedings of the 2000 symposium on Eye tracking research & applications (pp. 71-78). ACM.
12. SCHMIDT, R. W. (ED.). 1995. Attention and awareness in foreign language learning (Vol. 9). Natl Foreign Lg Resource Ctr.
13. SHUTE, V. J., VENTURA, M., & KIM, Y. J. 2013. Assessment and learning of qualitative physics in newton's playground. The Journal of Educational Research, 106(6), 423-430.
14. SHUTE, V., & VENTURA, M. 2013. Measuring and supporting learning in games: Stealth assessment
15. SIBERT, L. E., & JACOB, R. J. 2000, April. Evaluation of eye gaze interaction. In Proceedings of the SIGCHI conference on Human Factors in Computing Systems (pp. 281-288). ACM.
16. ZHU, Z., & JI, Q. 2004. Eye and gaze tracking for interactive graphic display. Machine Vision and Applications, 15(3), 139-148.