Computer Science and Information Systems 18(4):1287–1309 https://doi.org/10.2298/CSIS201201035N

Assessing Learning Styles through Eye Tracking for E-Learning Applications

Nahumi Nugrahaningsih[1], Marco Porta[2], and Aleksandra Klašnja-Milićević[3]

[1] University of Palangkaraya, Department of Informatics,
Kampus Unpar Tunjung Nyaho, Jl. Yos Sudarso , Palangkaraya 73112, Indonesia
nahumi@it.upr.ac.id
[2] University of Pavia, Department of Electrical, Computer and Biomedical Engineering,
Via A. Ferrata 5, 27100, Pavia, Italy
marco.porta@unipv.it
[3] University of Novi Sad, Department of Mathematics and Informatics,
Trg Dositeja Obradovica 3 – 21 000 Novi Sad, Serbia
akm@dmi.uns.ac.rs

Abstract. Adapting the presentation of learning material to the specific student's characteristics is useful to improve the overall learning experience and *learning styles* can play an important role to this purpose. In this paper, we investigate the possibility to distinguish between Visual and Verbal learning styles from gaze data. In an experiment involving first year students of an engineering faculty, content regarding the basics of programming was presented in both text and graphic form, and participants' gaze data was recorded by means of an eye tracker. Three metrics were selected to characterize the user's gaze behavior, namely, percentage of fixation duration, percentage of fixations, and average fixation duration. Percentages were calculated on ten intervals into which each participant's interaction time was subdivided, and this allowed us to perform time-based assessments. The obtained results showed a significant relation between gaze data and Visual/Verbal learning styles for an information arrangement where the same concept is presented in graphical format on the left and in text format on the right. We think that this study can provide a useful contribution to learning styles research carried out exploiting eye tracking technology, as it is characterized by unique traits that cannot be found in similar investigations.

Keywords: e-learning, learning models, learning styles, eye tracking, gaze behavior.

1. Introduction

The recent years, and especially the recent months, have seen a significant increase of e-learning solutions. However, in most cases the teaching material is the same for all students, without any distinction based on their specific "needs". *Learning styles* are a way to potentially identify how people learn best. Assessing learning styles to present the right material in proper ways to the user/learner can be of paramount importance in e-learning [1], [2].

It is a fact that there are different ways to learn, and different students will favor the learning modalities that are more suitable for them. Several investigations have highlighted the existence of bipolar learning styles, depending on whether, for example, a person prefers to learn by seeing or hearing, reflecting, or acting, reasoning in a logical or intuitive manner, visualizing or building mathematical models [3][31]. In general, researchers agree on the fact that learning materials should reflect students' learning styles. The huge amount of teaching resources currently offered in electronic form should therefore be adapted to the specific skills of the individual learner, in order to maximize the learning experience and improve learner achievements [4][32].

The most common way to assess learning styles is by means of questionnaires, through which students are asked to answer some questions aimed at discovering their preferred ways of learning. However, this kind of explicit assessment has some drawbacks ─ for instance, it may be considered long and boring, which causes careless responding, and the provided answers may not be sufficiently reliable. Thus, is it possible to automatically evaluate a person's learning style from the way he or she looks at the learning material? The aim of the paper is to identify the possibilities of *Eye tracking* technology to provide this kind of information, making learning style assessment a seamless procedure integrated into e-learning platforms ─ for example, through the analysis of the user's gaze behavior in the very initial stages of an e-learning course.

Incorporating eye tracking into adaptive e-learning systems by using data about pupil and gaze to indicate attentional focus and cognitive load levels can be useful in a process of adaptation to the requirements and needs of the learner. Personalization of an e-learning program based on the learner's cognitive load levels and learning styles calculated from eye-tracking data will impart the advantage of having a personal tutoring system into a wideband environment, with successful training by increasing information transfer and maintenance.

This is now realistic, as recent technological advances have enabled the development of affordable, robust, and mainstream eye-tracking solutions. Eye tracking is the process through which devices called *eye trackers* can detect the user's gaze direction [5]. In other words, an eye tracker identifies where a person is looking at (typically on a screen) and records the related gaze coordinates. Eye movements are characterized by very fast *saccades*, generally lasting less than 100 ms, interspersed with relatively steady periods of *fixations*, normally lasting between 100 and 600 ms. The main purpose of eye movements is to reallocate the gaze on the specific target, so that it can be clearly sensed on the fovea, the most sensitive area of the retina.

Eye trackers are becoming increasingly widespread nowadays, thanks to the availability of cheap devices. Current eye trackers are also non-invasive tools that do not constrain the user and allow to gather meaningful information in relatively simple ways. In the study presented in this paper, we exploited eye tracking technology to assess the user's Visual/Verbal learning styles from the way some slides presenting basic computer science topics (on the notion of *variable*, the concept of *algorithm*, and the *sequence*, *selection*, and *iteration* basic imperative programming constructs) were read/observed by 90 first-year engineering students.

Our study has two research objectives, one primary and one secondary. The primary research objective is: *Is it possible to distinguish Visual and Verbal learners from their gaze data recorded by an eye tracker?* Three metrics were selected to characterize the user's gaze behavior, namely, percentage of fixation duration, percentage of fixations, and average fixation duration. Our experiments were designed around this main purpose, through the presentation of basic computer science concepts in both textual and graphical form. However, as a secondary research objective, we also considered the possibility to *recognize the Active/Reflective*, *Sensible/Intuitive*, and *Sequential/Global* bipolar styles from learners' gaze behavior.

Gaze data were coupled with the outcomes of the Index of Learning Styles (ILS) questionnaire [6], one of the most widespread methods to evaluate users' learning styles. Even if a clear connection between the Visual/Verbal learning style could be found only for a specific information layout, we believe that our investigation can provide a constructive contribution to the field of e-learning in general, and to the area of automatic learning style assessment specifically. Exploiting eye tracking in this field is of paramount importance because it can potentially enable "intelligent" e-learning systems in which learning styles are assessed in a seamless way.

The paper is structured as follows. Section 2 presents a short summary of works that have exploited eye tracking technology for learning style assessment. Section 3 explains the main research questions at the basis of our study. Section 4 describes the methodology used for our investigation. Section 5 illustrates the performed analysis and the obtained results, which are then discussed in Section 6. Lastly, Section 7 outlines the conclusion and future work on the presented topic.

2. Background

This section provides an overview of eye tracking studies aimed at detecting learning and cognitive styles. A summary of the collected works is shown in Table 1. As can be seen, most investigations are focused on the Visual/Verbal learning styles. The order of the presented works is chronological.

Hughes et. al. [7] conducted an eye tracking study with 12 participants to investigate the difference between Verbalizer/Visualizer learners. The learning style was measured using the Verbalizer Visualizer Questionnaire (VVQ) by Kirby et al. [8]. Gaze data was recorded with an ASL 504 eye tracker. Stimuli were organized into ten "screens", each containing 20-25 video segments. The task for participants was to find a video that matched with a given topic. To avoid learning effects, the positions of text and visual components in the slides were alternately on the left and on the right. Since the VVQ results showed that no participants were in the verbalize group, the comparison was made only on visualizer and balanced learners. The eye features used were the average duration of fixations on slides, average fixation count, and average fixation duration. The statistical analysis revealed that there was a significant difference between the two groups. In particular, it indicated that balanced learners spent more time in the text area than visualizer learners. Despite variations in the layout, the statistical analysis revealed that participants' first fixations tended to be on the left side of the slide, regardless of the specific content in that area.

Tsianos at al. [9] tried to distinguish subjects into the wholist/analyst and verbal/imagery groups according to the Riding and Cheema's Cognitive Style Analysis (CSA) [10]. Twenty-one participants were involved in the experiment. The stimuli were web pages containing basic programming theories. The employed eye features were the ratio of fixation duration (i.e., the ratio between the times spent within image and text areas), the number of fixations on the page menu, and the experiment duration. The results of the statistical analysis on fixation ratios showed that imagers focused more on images, verbalizers more on text, and intermediates on both kinds of stimuli. The analysis on the number of fixations on the menu indicated that there was no difference among groups. Regarding session duration, imagers and intermediates devoted about the same time to read the whole content, while verbalizers spent considerably less time.

Al-Wabil et al. [11] observed the difference between visual and verbal learners according to Felder and Silverman's learning style [3]. Eight participants were involved in the study. The stimuli were six slides containing an introduction to statistics. The eye features employed in the study were total fixation duration, mean of fixation duration, and number of fixations. Eye features were compared without performing a statistical significance test. Acquired data showed that visual learners looked more at the multimedia area, while verbal learners looked more at the text area. Regarding the mean of fixation duration, there was no difference between visual and verbal learners. The comparison of the number of fixations showed that there was no difference, as all participants tended to have more fixations on text.

Mehigan et al. [12] tried to distinguish between visual and verbal learners, who were evaluated with an online survey [6] implementing the Felder and Silverman's learning style model. Several candidate participants were analyzed until a minimum of five visual learners and five verbal learners were found. The stimuli were composed of two slides: the first contained material about server-side programming, while the second contained a multiple-choice question to test the participants' comprehension level. The first slide was divided into two equal areas containing an image and text. Fixation count, total fixation time on the text area, and total fixation time on the image area were analyzed. The data showed that visual learners made more fixations on the graphic slide area than verbal learners. However, no statistical significance assessment was conducted to confirm this result. Regarding the fixation time in image and text areas, a visual inspection on correlation distribution revealed that students with longer fixation duration on visual content tended to be more visual in their learning style, while learners with longer fixation duration on textual content tended to be verbal.

Cao and Nishihara [13] conducted an eye tracking experiment with 38 participants. The main focus of their research was to find the difference between Visual/Verbal and sequential/global learners according to the Felder and Silverman learning styles. The stimuli were 11 slides through which the participants could freely navigate. To distinguish visual and verbal participants, fixation time was employed as a feature. The obtained results showed that even though visual participants spent more time on picture areas than verbal participants, the difference was not significant. The same trend also appeared in the text area. To discriminate between sequential and global learners, the features employed were fixation duration, saccadic length, and saccadic orientation (i.e., the angle between the horizontal line and the saccade direction line). Results showed that global learners tended to have shorter fixation durations and moved the eyes faster and with larger degrees. However, differences were not significant in this case either.

Alyahya [14] examined the different performance between verbal/visual students when they were observing a historical map. The experiment involved 62 female students and learning styles of participants were self-assessed with the Verbal-Visual Learning

Style Rating instrument [15]. Stimuli of the eye tracking experiment were derived from Minard’s map, (i.e., a graph that combines a geographic map with a bar graph, time series, and text to present the journey of Napoleon’s march from France to Moscow in 1812). After presenting content slides to participants, their comprehension was tested with 20 textual multiple-choice questions (which the author called a *verbal test*) and a *visual test* to verify how much maps and cities were recalled. The results of an ANOVA analysis showed that there was a significant variability in the results of the visual test among the participants with different learning style. The posthoc analysis results indicated that the visual group performed better than the *mostly visual* group. However, there was no significant difference for the verbal test. From a visual inspection of the accumulative heatmap of each learning style group, it was found that both groups spent about the same time on the text area. However, the difference was especially evident in the map area where visual learners watched more than verbal learners.

Nisiforou and Laghos [16] investigated the relation between eye movements and cognitive style. A total of 54 students participated in the experiment. The cognitive style of participants was evaluated with a paper-based Hidden Figures Test [17]. Based on the results of the evaluation, participants were grouped into Field Dependent (FD), Field Independent (FI), and Field Neutral (FN) participants. In an eye tracking experiment, participants were asked to answer four questions that were inspired by the Hidden Figure Test. Participants had to click on the shapes that were hidden in pictures. The results obtained from a visual inspection of the gazeplot indicated that FD participants had more random gazeplot compared to FI participants, who had more oriented and organized gazeplots. Moreover, one-way ANOVA analyses carried out on fixation count and saccade count showed that there were significant differences among groups.

Goswami et al. [18] observed the gaze behavior of 13 participants with different learning styles when they tried to identify errors in a project document. Learning styles were assessed using the Felder and Silverman Index of Learning Style. Stimuli were 14 pages containing 14 errors, which were marked as Areas of Interest (AOI). It is to be stressed that the purpose of this work was not to recognize the user’s learning styles (like in our study), but, instead, to compare the user’s performance according to learning styles. An evaluation was conducted to recognize effective (those who found more faults) and efficient (those who found faults faster) participants. The considered eye features were total fixation time per page, duration per page, linear saccade per page, total fixation per AOI, and duration per AOI. The results of multiple regression analyses indicated that total fixation, total fixation per AOI, and duration per AOI were factors that significantly contributed to achieve a high effectiveness. High effectiveness was shown by participants with sensible and sequential styles. As for efficiency, there were no factors that were positively significant; however intuitive and global participants tended to have an eye behavior that influenced efficiency in a negative way. The same negative tendency on effectiveness was also found on participants with a combination of verbal and linear styles.

Koc-Januchta et al. [19] carried out an investigation to explore the differences between visualizers and verbalizers according to how they look at pictures and text during the learning process. Through questionnaires, students were categorized based on their visual or verbal cognitive styles. Two different topics were used. The results showed that visualizers spent more time on images than verbalizers, and verbalizers spent more time reading text. Also, verbalizers observed non-informative picture areas earlier than visualizers. A similar study was carried out by Höffler et al. [20] (from the same research group) to validate the Object-Spatial Imagery and Verbal Questionnaire (OSIVQ) – which assumes a three-dimensional cognitive style model discriminating between object imagery, spatial imagery, and verbal dimensions. They found substantially different correlations of the different cognitive style scales with gaze behavior and visual-spatial ability. Participants scoring high on the object scale and/or the spatial scale of OSIVQ relied more heavily on pictures than on texts (indicated by high positive correlations with a joint gaze behavior score), while participants scoring high on the verbal scale tended to rely on texts (indicated by a negative, non-significant correlation). Additionally, only participants scoring high on the spatial scale tended to additionally have a high visuo-spatial ability, as indicated by a significant positive correlation.

Raptis et al. [21] presented two studies based on a multifactorial model. In both, participants carried out visual tasks with different characteristics, and eye tracking analysis discovered significant differences among participants characterized by different cognitive styles. In particular, the authors considered the Field Dependence-Independence (FD-I) cognitive style theory: while field-dependent users tend to prefer holistic ways for processing visual information, field-independent users tend to favor more analytical information processing approaches. The study revealed that the first category of users followed a more disoriented approach when performing visual search tasks, while the second category adopted a more organized visual strategy. Such differences suggested classification experiments in which different classifiers were trained with eye tracking data to infer the category a user belongs to.

Alhasan et al. [22] conducted a preliminary eye tracking study to analyze the pattern of learner behavior in order to obtain their learning style as a personalization aspect in an e-learning system. The electroencephalography (EEG) Emotive Epoc device was used to disclose learners with more accurate data. A method was developed to determine whether the verbal and visual learning styles reflect actual preferences in an e-learning environment based on the Felder and Silverman Learning Style Model. "Emotions" were exploited to exclude the periods of time when the learner was not focusing on learning. The primary experiment designed to test the combination of eye tracking and EEG confirmed operability and efficiency of this approach for studying and analyzing learning styles.

The studies that mainly guided the methodological choices of our investigation were [7], [9], [11], [12], [13], [14], and [19]. All of them have, as a main purpose, the recognition of learning styles from users' gaze behavior. They also include the Visual/Verbal styles, which are the primary research objective of this study. Moreover, our investigation introduces novel elements compared to these previous works, such as the fact that participants had no time limits. This choice allowed us to carry out an experiment closer to real learning scenarios, without sacrificing a time-dependent analysis. As will be illustrated in Section 5.2 (*Statistical Procedures*), such analysis was implemented through the subdivision of the single participants' interaction times with the learning stimuli into intervals, which is also an original aspect of our work.

Table 1. Studies using eye tracking for learning style detection.

Studies	Number of Participants	Learning Style Instruments	Learning Styles	Eye Features
[28]	22	Felder and Soloman	Visual/Verbal	Gaze paths, fixation count, fixation duration and average time for each fixation
[29]	28	Felder and Silverman Learning Style Model (FSLSM)	Visual/Verbal	Fixation duration, fixation count and the average time on each fixation
[30]	7	Felder-Silverman Index of Learning Styles	Visual/Verbal	The time that the participants gazed at text-based or graphic-based learning objects
[22]	48	Felder and Silverman Learning Style Model [3]	Visual/Verbal	Fixation count, average fixation duration
[21]	36	Field Dependence-Independence theory	Field dependent/field independent	Fixation count, fixation duration, saccade length, combined metrics
[20]	32	Object-Spatial Imagery and Verbal Questionnaire (OSIVQ, [23])	Object visualizers, spatial visualizers, and verbalizers	Dwell time (sum of durations from all fixations and saccades that hit the AOI in seconds) and revisits (number of returns to the AOI after the first visit)
[19]	32	Santa Barbara Learning Style Questionnaire (SBCSQ, [15]), Individual Differences Questionnaire [24], Vividness of Visual Imagery Questionnaire (VVIQ, [25]), Verbalizer – Visualizer Questionnaire [26],	Visual/Verbal	First gaze time (duration from start of the trial to the first hit of the AOI), dwell time (sum of durations of all fixations and saccades that hit the AOI), and transitions (movements from one AOI to another)

		Object-Spatial Imagery and Verbal Questionnaire (OSIVQ, shortened version, [23])		
[18]	13	Felder and Silverman Index of Learning Style [3]	Active/reflective, Sensible/intuitive, Visual/Verbal, Sequence/global	Total fixation time per page, duration per page, linear saccade per page, total fixation per AOI, duration per AOI
[16]	54	Hidden Figures Test [17]	Field dependent / field neutral/field independent	Fixation count, saccade count, average fixation duration, average saccade duration
[14]	62	Verbal-Visual Learning Style Rating [15]	Visual/Verbal	Total fixation duration on the text and map
[13]	38	Felder and Silverman Index of Learning Style [3]	Visual/Verbal, Sequential/global	Fixation duration, saccadic length, and the saccadic orientation (i.e. the angle between the horizontal line and the saccade line)
[12] *	10	Felder-Solomon Index of Learning Styles [6]	Visual/Verbal	Fixation count, total fixation time on text area, total fixation time on image area
[11] *	8	Felder and Silverman Index of Learning Style [3]	Visual/Verbal	Total fixation duration, average fixation duration, and fixation count
[9]	21	Riding and Cheema's Cognitive Style Analysis [10]	Wholist/analyst Verbal/imagery	Ratio of fixation duration (i.e. ratio between time spent in image and text areas), fixation count on the menu, and duration of the sessions
[7]	12	Kirby, Moore and Schofield's Verbalizer Visualizer Questionnaire [8]	Verbalizer/visualizer	Average duration of slide, average fixation count, average fixation duration

* Direct comparisons were performed without a statistical significance test

3. Research Questions

As already stated in the Introduction, the present study was based on a primary (RQ1) and on a secondary (RQ2) research questions. The reason for such a distinction is because our experiments were mainly designed to answer RQ1. Nevertheless, we wanted to verify whether, using the same data gathered for RQ1, it was also possible to answer RQ2.

The two research questions were:

RQ1. Is it possible to distinguish Visual and Verbal learners from the features of their gaze behavior (percentage of fixation duration, percentage of fixations, and average fixation duration) recorded by an eye tracker?

RQ2. Is it possible to recognize Active/Reflective, Sensible/Intuitive, and Sequential/Global learners from the features of their gaze behavior (percentage of fixation duration, percentage of fixations, and average fixation duration recorded by an eye tracker)?

4. Methodology

4.1. Participants

In total, 90 volunteer students participated in the experiment (57 males and 33 females, 18 years old on average). All of them were freshman Computer Engineering students of the Informatics Department of the University of Palangkaraya and had not attended any computer programming course yet. The recruitment occurred through announcements in bulletin boards in the department. All the participants, generally curious about eye tracking technology, were fully informed about the experiment procedures before starting them. No personal data were stored, as all the participants in the experiment were anonymously identified through numbers (only needed to match questionnaire data with eye tracking data). The participants did not get any academic credits for participating in the experiments, but they simply received their "gazeplots" (graphical representations indicating the visual scanpaths of their gaze) as "souvenirs".

4.2. Materials

To record gaze data, we employed the low-cost Eye Tribe ET-1000 eye tracker [27], with 60 Hz data sampling rate. Stimuli were displayed on a 21.5" monitor.

4.3. Procedure

The experiments were subdivided into two phases, namely *Experimental Phase 1* and *Experimental Phase 2*.

Experimental Phase 1. To preliminarily investigate their learning styles through a "traditional" approach, the participants were initially asked to complete the Index of Learning Styles (ILS) questionnaire [6]. The ILS questionnaire is an instrument composed of 44 multiple-choice questions which aims to distinguish four bipolar styles, namely Active/Reflective (AR), Sensible/Intuitive (SI), Visual/Verbal (VV), and Sequential/Global (SG). There are two answers (*a* and *b*) for each question. In our study, the original questionnaire was translated into Indonesian.

The Index of Learning Styles of each participant was calculated using the scoring sheet shown in Figure 1.

ACT/REF			SNS/INT			VIS/VRB			SEQ/GLO		
Q	a	b	Q	a	b	Q	a	b	Q	a	b
1	___	___	2	___	___	3	___	___	4	___	___
5	___	___	6	___	___	7	___	___	8	___	___
9	___	___	10	___	___	11	___	___	12	___	___
13	___	___	14	___	___	15	___	___	16	___	___
17	___	___	18	___	___	19	___	___	20	___	___
21	___	___	22	___	___	23	___	___	24	___	___
25	___	___	26	___	___	27	___	___	28	___	___
29	___	___	30	___	___	31	___	___	32	___	___
33	___	___	34	___	___	35	___	___	36	___	___
37	___	___	38	___	___	39	___	___	40	___	___
41	___	___	42	___	___	43	___	___	44	___	___

Total (sum X's in each column)

ACT/REF		SNS/INT		VIS/VRB		SEQ/GLO	
a	b	a	b	a	b	a	b
___	___	___	___	___	___	___	___

(Larger – Smaller) + Letter of Larger (see below*)

____	____	____	____

*Example: If you totaled 3 for a and 8 for b, you would enter 5b in the space below.

Fig. 1. ILS Scoring sheet

The result score is an odd number between 1 and 11, whose interpretation, according to Felder and Soloman, is as follows:

- If the score is 1 or 3: the respondent is fairly well balanced on the two dimensions of that scale.
- If the score is 5 or 7: the respondent has a moderate preference for one dimension of the scale and will learn more easily in a teaching environment which favours that dimension.
- If the score is 9 or 11, the respondent has a very strong preference for one dimension of the scale and may have real difficulties when learning in an environment which does not support that preference.

Experimental Phase 2. Subsequently, after three days from the Experimental Phase 1, the participants also attended an eye tracking experiment. The participants were not informed that this trial was connected with the questionnaire they had answered in Phase

1. A within-subjects experimental design was used, in which participants tried all the available conditions.

The eye tracking experiment was conducted in a quiet room, with artificial illumination from the ceiling. The participant in the test was seated at about 55 cm from the monitor. The task was to read and try to understand the topics presented in a group of slides. No time limit was set for each slide, so that the participants could learn at their own pace (a new slide was loaded by pressing the space bar).

In total, there were seven slides. The first one contained a description of the task; the second consisted of a graphical overview of the topics; the third explained the basic notion of *variable*; the fourth presented the concept of *algorithm*; and the fifth, sixth and seventh slides, respectively, covered the three basic imperative programming constructs, namely *sequence*, *selection*, and *iteration*.

In this study, we focused on slides from the 4th to the 7th in the above list, that, in the following, we will identify as slides *a*, *b*, *c*, and *d*. Figures 2a-2d show the translation of the original slides (written in Indonesian) into English.

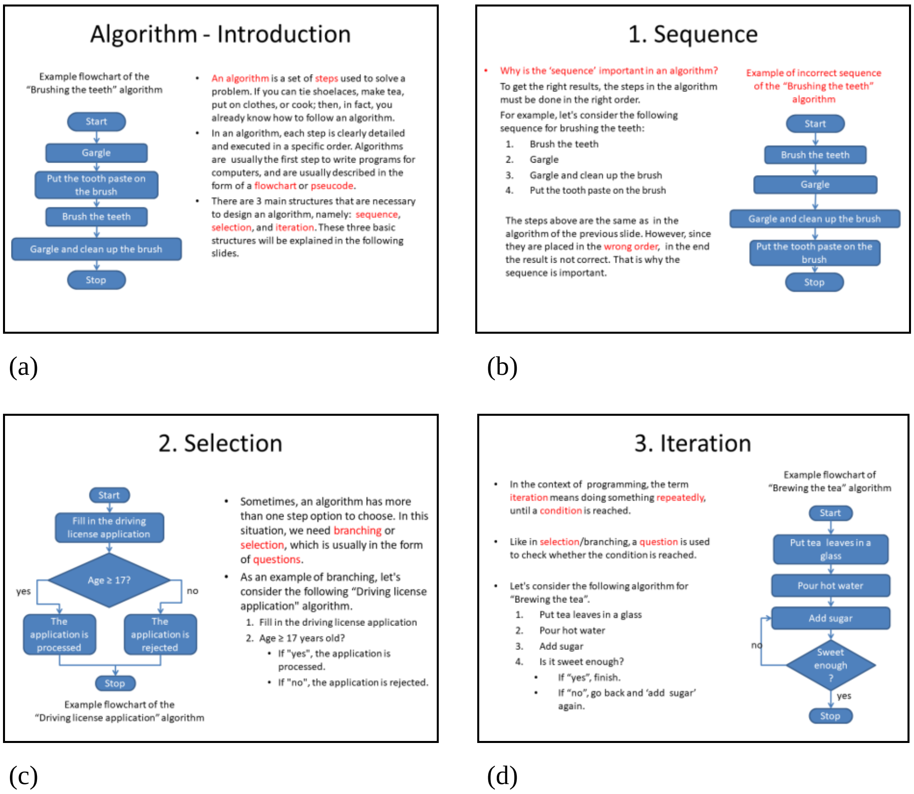

Fig. 2. English version of the slides used as stimuli in the eye tracking experiment

5. Analysis of Eye Tracking Data and Results

In each slide, we defined two AOIs (Areas Of Interest): one for the text section and another for the picture region (Figure 3). As can be seen from Figure 2, text and pictures were alternately on the left and on the right within slides.

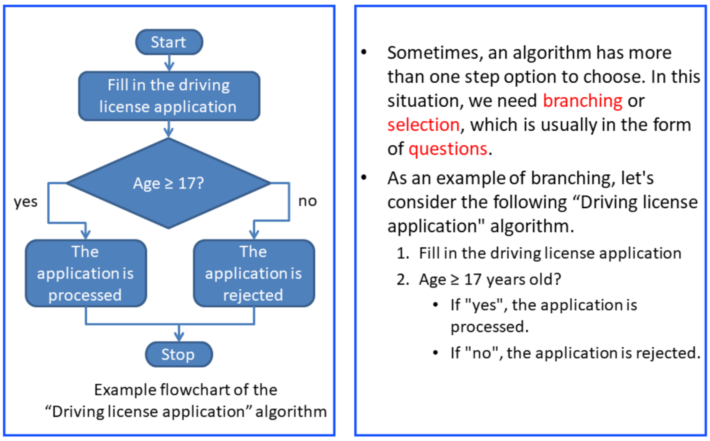

Fig. 3. Example of AOIs in a slide

The *independent* variables of the eye tracking study were the position of the picture and of the text areas on the slides (left-right or right-left).

The *controlled variables* were the textual and graphical contents displayed in the slides (arranged as shown in Figure 2).

The *dependent* variables, besides the questionnaire outcomes for Phase 1, in Phase 2 were the *percentage of fixation duration* (i.e., the percentage of fixation time on the AOI), the *percentage of fixations* (i.e., the percentage of fixations detected on the AOI), and the *average fixation duration*. Percentages were preferred to absolute values because the time spent on each slide by each participant was different.

For a temporal analysis of eye behavior, we subdivided the whole time spent by each participant on a slide into ten intervals. For each slide and each interval, we calculated the percentage of fixation duration (over the total time spent on the slide), the percentage of fixations (over the total number of fixations detected on the slide), and the average fixation duration up to that interval.

Unfortunately, six of the 90 participants did not fill in the questionnaire completely. Other four participants failed the eye tracking calibration procedure (consisting in fixating the center of a circle appearing in different positions of the screen). Moreover, 25 participants tried the test more than once, due to problems occurring in the data recording phase. Thus, in the end, we decided to consider only eye data from the surely reliable 55 participants.

5.1. Score Distributions

As shown by the histograms in Figure 4, the scores obtained from the Felder-Silverman questionnaire were not evenly distributed.

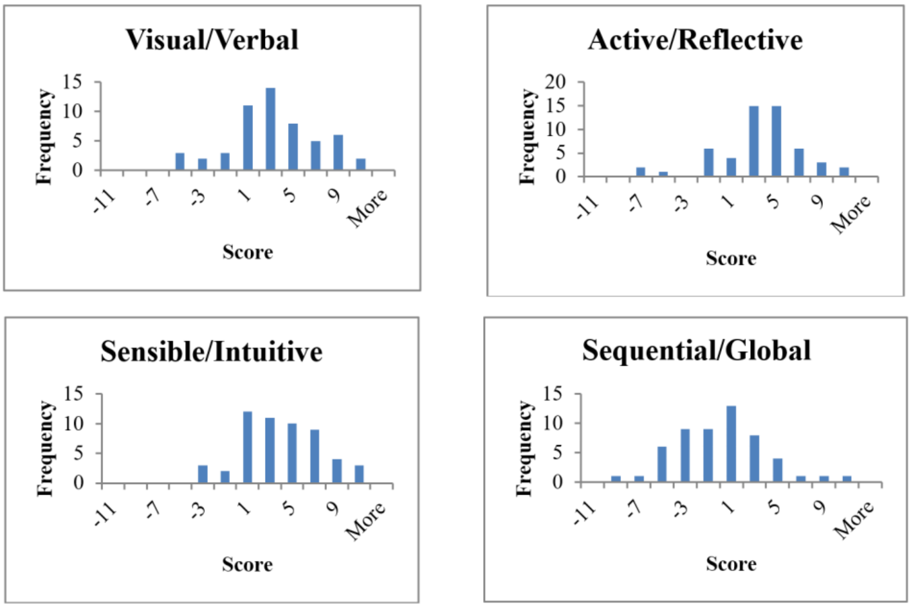

Fig. 4. Score histograms obtained from the answers to the Felder-Silverman questionnaire

For this reason, instead of classifying participants by the score threshold (as suggested by Felder and Soloman), we grouped them based on the median (MED) and median absolute deviation (MAD) values of the score. Specifically, we identified three groups:

- *Group 1*, with score < MED – MAD
- *Group 2*, with score > MED + MAD
- *Group 3*, with score in the range (MED – MAD) ÷ (MED + MAD)

Since learning styles are bipolar measurements, this classification can be interpreted as a learning style "tendency" of participants in the three groups. For example, for the Visual/Verbal case, Group 1 means "more verbal than visual", Group 2 "more visual than verbal", and Group 3 "between visual and verbal". Table 2 shows the number of participants in each group for the four kinds of learning styles.

Table 2. Studies using eye tracking for learning style detection

	Visual/Verbal MED = 3, MAD = 2	**Active/Reflective MED = 3, MAD = 2**	**Sensible/Intuitive MED = 3, MAD = 2**	**Sequential/Global MED = 1, MAD = 2**
Group 1	8	9	5	17
Group 2	13	11	16	7
Group 3	34	35	34	31

5.2. Statistical Procedure

For all the three selected metrics (percentage of fixation duration, percentage of fixations, and average fixation duration), we used the Shapiro-Wilk test to verify the normality of data distributions in the ten intervals. Since distributions were not normal in numerous cases, and various attempts to transform data using several functions were not successful, we carried out a non-parametric statistical analysis. We therefore considered medians instead of means.

The next step of the analysis had the purpose to verify whether the specific slide influenced the three metrics. This was done by means of the Friedman's test applied to each learning style group (Group 1, Group 2, and Group 3) separately. The obtained results indicated that differences among slides were significant in most cases, and therefore it was not possible to consider the four slides together. Thus, we investigated whether a common behavior could be found considering the couples of slides with the same structure, (i.e., the two slides with the picture on the left and the text on the right slides *a* and *c* in Figure 2), and the two slides with the opposite arrangement (i.e., slides *b* and *d*).

We considered the four kinds of learning styles – Visual/Verbal (VV), Active/Reflective (AR), Sensible/Intuitive (SI), and Sequential/Global (SG) – separately, and the three metrics for each of them. In both the text and picture areas, in each of the six (3 groups x 2 areas) cases of each learning style and metric, we counted the number of occurrences in which the influence of the slide was not significant, with a 5% significance level. This value is traditionally and universally used in statistics as the significance level for decisions.

Although it was not possible to find cases in which the values of the metrics were independent of the slide in all 10 intervals for all three learning style groups, we considered as acceptable, or *valid*, those cases where the effect of the *slide* factor was not significant in at least seven intervals out of ten. For the two slides with the picture on the left and the two slides with the picture on the right, respectively, Tables 3 and 4 show these valid occurrences for each learning style category and metric, indicating whether they are related to the picture area (P), the text area (T), or both (PT).

As can be seen from Table 3, when the picture is on the left, it is never possible to consider the average fixation duration (no pair of bipolar styles has at least seven non-significant differences between slides *a* and *c*). The percentage of fixations is potentially useful for the VV, AR, and SG categories only on the text area. Lastly, the percentage of fixation duration is exploitable on both the picture and text areas for all learning style categories except SG (for which only the text region can be analyzed).

Table 3. Cases with at least seven non-significant differences between slides a and c (picture on the left and text on the right), for each learning style category and metric (P = picture area, T = text area)

	VV	**AR**	**SI**	**SG**
%FixDur	PT	PT	PT	T
%Fix	T	T		T
AvgFixDur				

Table 4. Cases with at least seven non-significant differences between slides b and d (picture on the right and text on the left), for each learning style category and metric (P = picture area, T = text area)

	VV	**AR**	**SI**	**SG**
%FixDur	T			
%Fix	P	PT	PT	PT
AvgFixDur	P	P		PT

When the picture is on the right (Table 4), the percentage of fixation duration can be considered only for the VV category and on the text area. The percentage of fixations is potentially useful on both the picture and text regions for all learning styles, except for VV (for which only the picture area can be studied). In regard to the average fixation duration, it can be used on both the text and the picture areas for SG, and only on the picture area for VV and AR.

After identifying the valid cases for metrics, learning style categories, and slide regions, the last step was using the Kruskal-Wallis test to find possible connections (i.e., relationships) between the metrics' values and learning style groups (Group 1, Group 2, and Group 3). Slides *a* and *c* (picture on the left) and slides *b* and *d* (picture on the right) were considered distinctly.

5.3. Results

For each metric, learning style category, AOI, and interval, we searched for valid cases with significant relations (5% significance level) between metric value and learning style group. This happened in very few occurrences, as shown in Tables 5 and 6.

As can be seen, the only metric with significant relations in both slides was, for slides a and c (i.e., picture on the left and text on the right), the percentage of fixation duration in the text area, for the Visual/Verbal style and in intervals 9 and 10. Hence, according to our analysis, in slides having a picture on the left and a corresponding text description on the right, the percentage of fixation duration up to the last part of the interaction (intervals 9 and 10), can be exploited to distinguish the groups of Visual/Verbal learners.

Table 5. Picture on the left and text on the right

Slide a:	*Slide c:*
• Percentage of fixation duration, VV, text area: in interval 9 ($\chi^2(2) = 7.237$, $p = .027$) and in interval 10 ($\chi^2(2) = 8.306$, $p = .016$)	• Percentage of fixation duration, VV, text area: in interval 9 ($\chi^2(2) = 6.421$, $p = .04$) and in interval 10 ($\chi^2(2) = 8.092$, $p = .017$) • Percentage of fixation duration, SI, text area: in interval 9 ($\chi^2(2) = 6.709$, $p = .035$) and in interval 10 ($\chi^2(2) = 6.304$, $p = .043$)

Table 6. Picture on the right and text on the left

Slide b:	*Slide d*:
• Average fixation duration, AR, text area: in interval 1 ($\chi2(2) = 7.449$, $p = .024$) • Average fixation duration, AR, picture area: in interval 2 ($\chi2(2) = 7.217$, $p = .027$)	No significant instances

In particular, pairwise comparisons (carried out using the Dunn-Bonferroni test) allowed to determine that, for both slides and both intervals, the difference was significant for Group 1 and Group 2, with the percentage of fixation duration always higher for Group 1. This means that, considering at least the first 90% of the interaction time with the slide, the text area was observed more than the picture region by Verbal learners and less by Visual learners. This is also evident from Figure 5, which shows the evolution over time (medians calculated up to each interval) of the percentage of fixation duration for the VV learning style in the text area in slides *a* (left) and *c* (right). Figure 6 shows the box plots indicating the values of the medians of each group for each slide (*a* and *b*) and intervals 9 and 10.

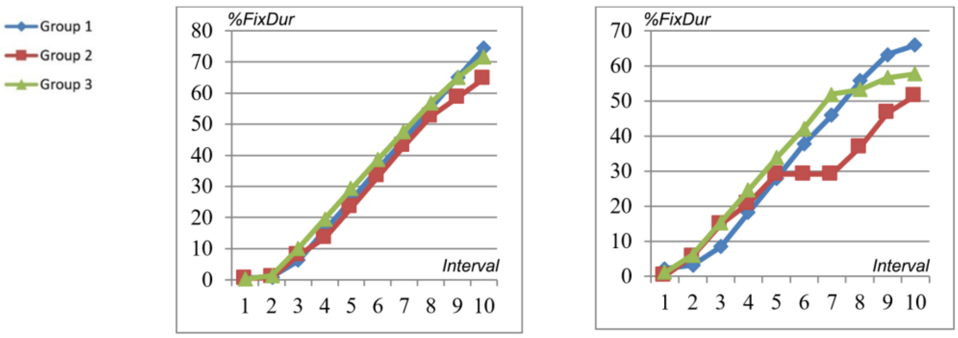

Fig. 5. Evolution of the percentage of fixation duration on the text area for the Visual/Verbal learning styles and the three learner groups in slide *a* (left) and in slide *c* (right)

In a boxplot, the bottom and top of the box indicate the 25th and 75th percentiles (i.e., the percentages of fixation duration corresponding to, respectively, the 25% and the 75% of the gathered data), while the inner band designates the 50th percentile (i.e., the medians); the ends of the whiskers represent the smallest and largest non-outlier values; circles denote outliers standing more than 1.5 box-lengths above or below the box; and stars indicate extreme values, standing more than three box-lengths above or below the box.

6. Discussion

Eye-tracking technology can be useful for implicitly classifying users based on their high-level cognitive processes (i.e., cognitive styles) in real-time while performing activities with varying characteristics (e.g., type complexity). In the study presented in this paper, we exploited eye tracking technology to assess the user's Visual/Verbal learning styles from the way some slides presenting basic computer science topics (on the notion of *variable*, the concept of *algorithm*, and the *sequence*, *selection*, and *iteration* basic imperative programming constructs).

Three metrics were selected to characterize the user's gaze behavior, namely, percentage of fixation duration, percentage of fixations, and average fixation duration. Our experiments were designed around this main purpose, through the presentation of basic computer science concepts in both textual and graphical form. However, as a secondary research objective, we also considered the possibility to *recognize the Active/Reflective, Sensible/Intuitive,* and *Sequential/Global* bipolar styles from learners' gaze behavior.

Gaze data were coupled with the outcomes of the Index of Learning Styles (ILS) questionnaire [6], one of the most widespread methods to evaluate users' learning styles. A connection between the Visual/Verbal learning styles was found for a specific information layout, which gives a constructive contribution to the field of e-learning in general, and to the area of automatic learning style assessment specifically. Exploiting eye tracking in this field is of paramount importance because it can enable "intelligent" e-learning systems in which learning styles are assessed in a seamless way.

According to our results, the answer to the primary research question of our study (i.e., "Is it possible to distinguish Visual and Verbal learners from the features of their gaze behavior – percentage of fixation duration, percentage of fixations, and average fixation duration – recorded by an eye tracker?") is partially positive:

- A relation between gaze behavior and learners' group (groups obtained from our modified interpretation of the Felder-Silverman questionnaire outcomes, as illustrated in sub-section 5.1, Score Distributions) could be found only for Group 1 (participants who were classified as more verbal than visual) and Group 2 (participants who were classified as more visual than verbal), but not for Group 3 (participants who were classified as being between visual and verbal), which was the largest.
- The relation between gaze behavior and Groups 1 and 2 could be found only for slides having the picture on the left and the text description on the right, not for the opposite case.

Specifically, the percentage of fixation duration on the text area, computed up to intervals 9 and 10 (i.e., up to the last part of the slide reading/observation process), gives clear information about the user's style group (Group 1 or Group 2). This indicates that, if most of the time (at least 90%) spent on the slide is evaluated, the Visual/Verbal learner can be successfully recognized.

As regards the secondary research question of our study, (i.e., "Is it possible to recognize AR, SI, and SG learners from the features of their gaze behavior – percentage of fixation duration, percentage of fixations, and average fixation duration – recorded by an eye tracker?"), the answer is negative: for no metric, significant relations with the three learners' groups (Group 1, Group 2, and Group 3) could be found. This, however, was partially expected, as our experiments were specifically designed with the first research question in mind. Indeed, also looking at the literature, eye tracking has been rarely used in studies aimed at recognizing styles other than Visual and Verbal.

Due to the peculiarity of the experiments, we have implemented in our study, a direct comparison with previous works is not possible. The novelty of our approach is due to three main factors:

- The subdivision of participants into three groups, based on the outcomes of the Felder-Silverman questionnaire, using MED and MAD to define score intervals;
- The absence of time limits for participants while reading or observing the content of the presented slides, to make the experiment more similar to real learning scenarios;
- A temporal analysis carried out by subdividing the time taken by each participant to read/observe the content of each slide into (ten) intervals.

A limitation of our study is the simple structure of slides, which may prevent our results to be generalized to more complex layouts. Moreover, the subjects of the slides (notion of variable, concept of algorithm, and the three basic imperative programming constructs) are very specific, and this may have influenced the results. Also, all the participants were about the same age (18) and engineering students.

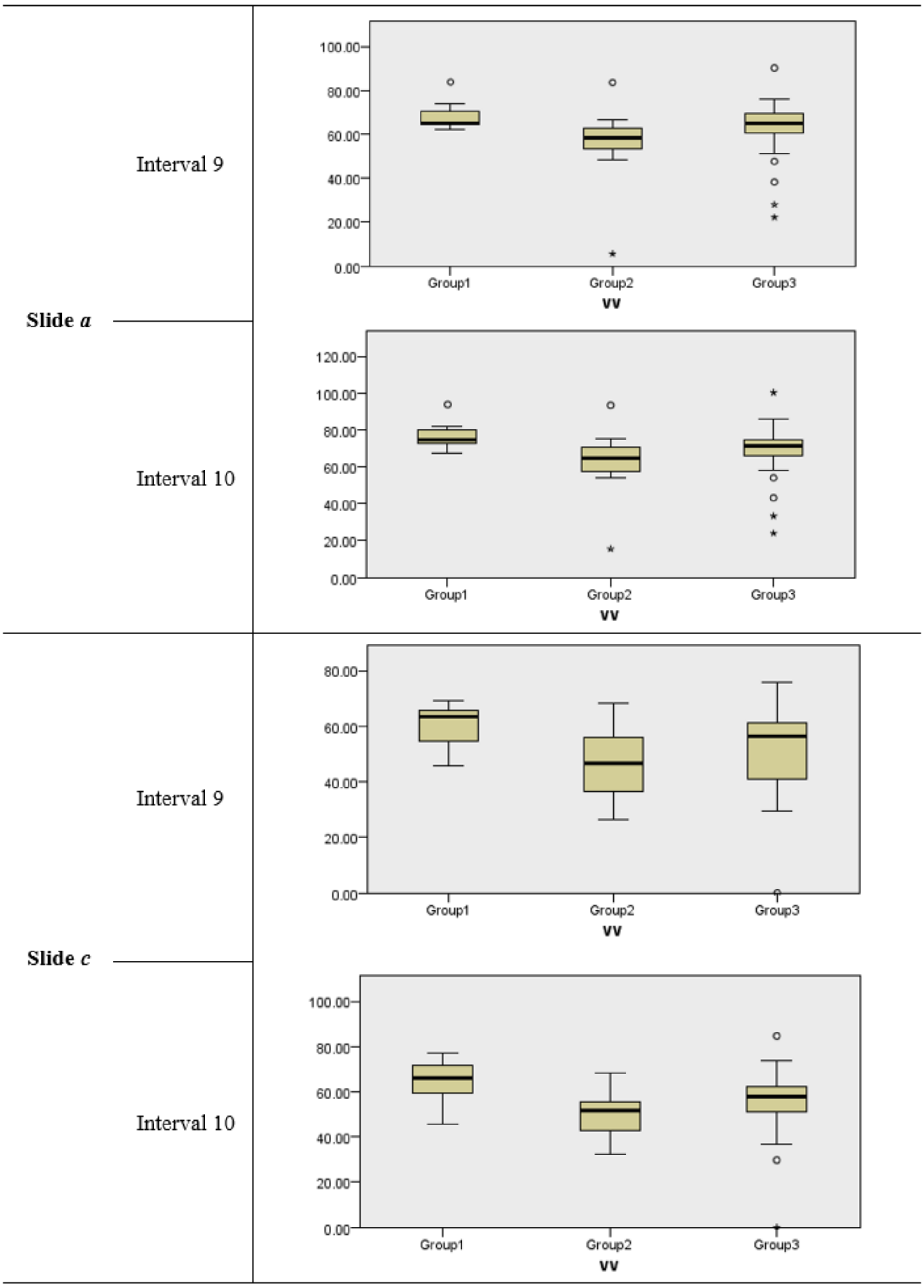

Fig. 6. Boxplots for each group, interval, and slide (Visual/Verbal learning styles)

7. Conclusions

In this paper, we have studied the possibility to recognize learning styles from the way users look at learning material, focusing in particular on the Visual/Verbal case. The content was basically structured into a two-column layout, with either an image on the left and text on the right or vice versa. The Index of Learning Styles questionnaire was exploited to preliminarily assess the styles of the participants in the experiments, to find possible connections between their gaze behavior and potential associated styles. The participants were grouped based on median and median absolute deviation of the scores obtained from the questionnaire. For a given bipolar learning style category, three groups were created which included participants who were “more towards one style” (Group 1), “more towards the other style” (Group 2), or “somewhere in the middle” (Group 3). This allowed us to deal with the unbalanced subdivisions of the participants in the two opposite learning style sets (such as Visual/Verbal).

Since gaze data distributions were not normal, the examination was carried out using non-parametric statistics. Three gaze metrics were considered, namely percentage of fixation duration, percentage of fixations, and average fixation duration. Percentage values allowed us to take into account the fact that the participants had no time limits and could read/observe a slide for how long they wanted. The time-dependent analysis was implemented through the subdivision of the whole interaction time with the slide into ten intervals.

Significant relations between the Visual/Verbal style and gaze behavior were found for the content layout in which the image is on the left and text is on the right. Specifically, clear distinctions between Groups 1 and 2 were identified using the percentage of fixation duration: considering at least the first 90% of the interaction time with the slide (i.e., measuring values of the metric up to intervals 9 or 10), the text region was looked at more than the picture area by verbal learners and less by visual learners.

Further research can continue to explore different design formats and deal with various types of illustrations, different difficulties of text and topics, and their impact on the learning styles of visualizers and verbalizers. It would have been useful to observe in detail how verbalizers learn only from text and how visualizers learn only from images.

We also tried to recognize other kinds of learning styles (Active/Reflective, Sensible/Intuitive, and Sequential/Global) using the same experimental material. However, as we could have expected, the results were not satisfying, because the investigation of these learning styles would have required different presentations of content, which we will consider in the future. Future work will also include further experiments with new topics, different content layouts, and more varied participants.

The automatic recognition of users’ learning styles is a very important step towards *intelligent adaptive learning platforms*. To achieve an adaptive e-learning system, it is essential to monitor the learner behavior dynamically to diagnose their learning style. Eye tracking can serve that purpose by investigating the eye gaze movement while engaging in the e-learning environment. It would be also useful to consider an application of eye tracking technology in combination with other biosensor systems. Additional tools and analytical data might explore hidden patterns in user behavior and activities. In particular, this should be taken into account when working on the implementation of adaptive tutoring systems..We think that the research presented in this paper can provide a useful contribution to gaze-based learning style research, stimulating further studies on the subject.

References

1. Jena, R. K.: Predicting students' learning style using learning analytics: a case study of business management students from India. Behaviour & Information Technology, Vol. 37, N. 10-11, 978-992. (2018)
2. Heidrich, L., Barbosa, J.L.V., Cambruzzi, W., Rigo, S.J., Martins, M.G., Dos Santos, R.B.S.: Diagnosis of learner dropout based on learning styles for online distance learning. Telematics and Informatics, Vol. 35, Issue 6, 1593-1606. (2018)
3. Felder, R.M., Silverman, L.K.: Learning and Teaching Styles in Engineering Education. Engr. Education, 78(7), 674–681. (1988)
4. Franzoni, A.L., Assar, S.: Student Learning Styles Adaptation Method Based on Teaching Strategies and Electronic Media. Educational Technology and Society, 12(4), 15–29. (2009)
5. Duchowski, A.: Eye Tracking Methodology – Theory and Practice (2nd edition), Springer, London. (2007)
6. Felder, R.M., Soloman, B.A.: Index of Learning Styles (ILS). Available: http://www4.ncsu.edu/unity/lockers/users/f/felder/public/ILSpage.html [Accessed: 18 May 2020].
7. Hughes, A., Wilkens, T., Wildemuth, B.M., Marchionini, G.: Text or pictures? An eyetracking study of how people view digital video surrogates. In Proc. 2003 International Conference on Image and Video Retrieval. Springer, 271–280. (2003)
8. Kirby, J.R., Moore, P.J., Schofield, N.J.: Verbal and visual learning styles. Contemp. Educ. Psychol. 13, 169–184. (1988)
9. Tsianos, N., Germanakos, P., Lekkas, Z., Mourlas, C., Samaras, G.: Eye-tracking users' behavior in relation to cognitive style within an e-learning environment. In Proc. 9th IEEE International Conference on Advanced Learning Technologies (ICALT 2009), 329–333. (2009)
10. Riding, R., Cheema, I.: Cognitive styles—an overview and integration. Educ. Psychol. 11, 193–215. (1991)
11. Al-Wabil, A., ElGibreen, H., George, R.P., Al-Dosary, B: Exploring the validity of learning styles as personalization parameters in eLearning environments: An eyetracking study. In Proc. 2010 2nd International Conference on Computer Technology and Development (ICCTD), 174–178. (2010)
12. Mehigan, T.J., Barry, M., Kehoe, A., Pitt, I.: Using eye tracking technology to identify visual and verbal learners. In Proc. 2011 IEEE International Conference on Multimedia and Expo, 1–6. (2011)
13. Cao, J., Nishihara, A.: Understanding Learning Style by Eye Tracking in Slide Video Learning. J. Educ. Multimed. Hypermedia 21, 335–358. (2012)
14. Alyahya, S.M.: Minard's graph of Napoleon's March to Moscow: An eye tracking study of cognitive processing. Ph.D. Thesis (advisers: Linda L.L., Gall, J.E.), University of Northern Colorado. (2014)
15. Mayer, R.E., Massa, L.J.: Three facets of visual and verbal learners: Cognitive ability, cognitive style, and learning preference. J. Educ. Psychol. 95, 833. (2003)
16. Nisiforou, E., Laghos, A.: Field Dependence–Independence and Eye Movement Patterns: Investigating Users' Differences Through an Eye Tracking Study. Interact. Comput. 28, 407–420. (2016)
17. Ekstrom, R.B., French, J.W., Harman, H.H., Dermen, D.: Manual for kit of factor-referenced cognitive tests. Educational Testing Service. Available: https://www.ets.org/Media/Research/

pdf/Manual_for_Kit_of_Factor-Referenced_Cognitive_Tests.pdf [Accessed: 12 June 2020]. (1976)
18. Goswami, A., Walia, G., McCourt, M., Padmanabhan, G.: Using Eye Tracking to Investigate Reading Patterns and Learning Styles of Software Requirement Inspectors to Enhance Inspection Team Outcome. In Proc. 10th ACM/IEEE International Symposium on Empirical Software Engineering and Measurement, ESEM '16. ACM, New York, NY, USA, 34:1–34:10. (2016)
19. Koć-Januchta, M., Höffler, T., Thoma, G. B., Prechtl, H., Leutner, D.: Visualizers versus verbalizers: Effects of cognitive style on learning with texts and pictures–An eye-tracking study. Computers in Human Behavior, 68, 170-179. (2017)
20. Höffler, T. N., Koć-Januchta, M., and Leutner, D.: More Evidence for Three Types of Cognitive Style: Validating the Object-Spatial Imagery and Verbal Questionnaire Using Eye Tracking when Learning with Texts and Pictures. Applied cognitive psychology, 31(1), 109-115. (2017)
21. Raptis, G. E., Katsini, C., Belk, M., Fidas, C., Samaras, G., and Avouris, N.: Using eye gaze data and visual activities to infer human cognitive styles: method and feasibility studies. In Proceedings of the 25th Conference on User Modeling, Adaptation and Personalization, 164-173. (2017)
22. Alhasan, K., Chen, L., hen, F.: An Experimental Study of Learning Behavior in an ELearning Environment. In Proc. of the 16th IEEE International Conference on Smart City (SmartCity-2018). (2018)
23. Blazhenkova, O., Kozhevnikov, M.: The new object–spatial–verbal cognitive style model: Theory and measurement. Applied Cognitive Psychology, 23, 638–663. (2009)
24. Paivio, A., Harshman, R.: Factor analysis of a questionnaire on imagery and verbal habits and skills. Canadian Journal of Psychology/Revue canadienne de psychologie, 37(4), 461. (1983)
25. Marks, D.F.: Visual imagery differences in the recall of pictures. British Journal of Psychology, 64, 17-24. (1973)
26. Richardson, A.: Verbalizer-Visualizer: a cognitive style dimension. Journal of Mental Imagery, 1, 109-126. (1977)
27. The Eye Tribe. Available: https://en.wikipedia.org/wiki/The_Eye_Tribe [Accessed: 21 July 2020].
28. Luo, Z., O'Steen, B., & Brown, C. (2020). The use of eye-tracking technology to identify visualisers and verbalisers: accuracy and contributing factors. *Interactive Technology and Smart Education*.
29. Luo, Z., & Wang, Y. (2019). Eye-tracking technology in identifying visualizers and verbalizers: data on eye-movement differences and detection accuracy. *Data in brief*, 26, 104447.30. El Guabassi, I., Bousalem, Z., Al Achhab, M., & EL Mohajir, B. E. (2019). Identifying learning style through eye tracking technology in adaptive learning systems. *International Journal of Electrical & Computer Engineering (2088-8708)*, *9*(5).
31. Wibirama, S., Sidhawara, A. P., Pritalia, G. L., & Adji, T. B. (2020, September). A Survey of Learning Style Detection Method using Eye-Tracking and Machine Learning in Multimedia Learning. In *2020 International Symposium on Community-centric Systems (CcS)* (pp. 1-6). IEEE.
32. Alemdag, E., & Cagiltay, K. (2018). A systematic review of eye tracking research on multimedia learning. *Computers & Education*, *125*, 413-428.

Nahumi Nugrahaningsih is a researcher at the University of Palangkaraya (Indonesia). She obtained her master's degree in Informatics from the Bandung Institute of Technology (Indonesia) in 2007 and her PhD in Electronics, Computer Science and

Electrical Engineering from the University of Pavia (Italy) in 2017. Her research interests include eye tracking and vision-based user interfaces for e-learning and biometric applications.

Marco Porta is associate professor at the University of Pavia (Italy). He received the master's degree in Electronic Engineering from the Politechnic of Milan (Italy) in 1996 and the PhD degree in Electronic and Computer Engineering from the University of Pavia in 1999. His research interests include eye tracking, vision-based perceptive interfaces, e-learning, user centered design, and Human-Computer Interaction in general.

Aleksandra Klasnja-Milicevic holds the position of Associate Professor at the Faculty of Sciences, University of Novi Sad (Serbia), where she received her PhD in Computer science, in 2013. Her research interests include e-learning, technology-enhanced learning (TEL), adaptivity and personalization, educational technology, learning analytics, and recommender systems.

Received: December 20, 2019; Accepted: May 02, 2020