Understanding Two Graphical Visualizations from Observer's Pupillary Responses and Neural Network

Md Zakir Hossain[†]
RSCS, ANU, Canberra, Australia
zakir.hossain@anu.edu.au

Tom Gedeon
RSCS, ANU, Canberra, Australia
tom@cs.anu.edu.au

Atiqul Islam
RSCS, ANU, Canberra, Australia
atiqul.islam@anu.edu.au

ABSTRACT

This paper investigates observers' pupillary responses while they viewed two graphical visualizations (circular and organizational). The graphical visualizations are snapshots of the kind of data used in checking the degree of compliance with corporate governance best practice. Six very similar questions were asked from 24 observers for each visualization. In particular, we developed a neural network based classification model to understand these two visualizations from temporal features of observers' pupillary responses. We predicted that whether each observer is more accurate in understanding the two visualizations from their unconscious pupillary responses or conscious verbal responses, by answering relevant questions. We found that observers were physiologically 96.5% and 95.1% accurate, and verbally 80.6% and 81.3% accurate, for the circular and organizational visualizations, respectively.

CCS CONCEPTS

Human-centered computing → Visualization → Empirical studies in visualization → Visualization design and evaluation methods

KEYWORDS

Information Visualizations, Pupillary Responses, Neural Network

ACM Reference format:

Md Zakir Hossain, Tom Gedeon and Atiqul Islam. 2018. Understanding Two Graphical Visualizations from Observer's Pupillary Responses and Neural Network. In *Proceedings of the 30th Australian Computer-Human Interaction Conference (OzCHI '18), December 4–7, 2018, Melbourne, VIC, Australia* DOI: 10.1145/3292147.3292187, ISBN: 978-1-4503-6188-0/18/12

1 Introduction

The term 'Visualization' refers to the focused and systematic visual representation of information in the form of graphs, tables and diagrams. Information visualization provides understanding of datasets with insight and clarity [1]. It is traditionally seen as a set of methods for supporting humans to understand large and complex data sets in a simple and easy way; consider insights from user interactions in developing interactive visualization interfaces. Efficient visualization enables the understanding of data from a suitable visual representation. The differences in visual representations emphasize particular visualization goals or specific data trends. However, how people perceive and interact with a visualization tool can strongly influence their understanding of data as well as that system's usefulness [2]. To illustrate, visualization that does not support cognitive processing creates pressure on human working memory. Human working memory has limited capacity where the information is processed [3]. Thus an unsuitable information visualization system may put a high cognitive load on users. Again, the choice of best visualization technique varies from user to user. Therefore, user-adaptive visualizations improve user experience through addressing each user's ability and specific need [4, 5].

ACM ISBN 978-1-4503-6188-0/18/12…$15.00
https://doi.org/10.1145/3292147.3292187

When someone is paying attention by looking at something, then pupil size varies according to their attention to the task [6]. The autonomic nervous system regulates the changes in pupil size, which reflects cognitive load, attention, confusion and a range of emotions that are beyond a person's control [7]. Thus pupil changes provide valuable information about the internal cognitive state of the observer of visualization.

In a previous study, two visualizations were compared from observers' visual fixations and saccades [8] where circular was referred to as radial and organizational was referred to as hierarchical due to an incorrect source. The correct sources are [13, 14]. Two similar visualizations were compared in that study showing that it is possible to differentiate between two visualizations using simple eye gaze path metrics. In this study, we focus on the temporal features of the observers' eye gaze data. We develop a neural network model for the classification performance using extracted features from observers' pupillary responses. We focus on different neural learning algorithms and the number of hidden nodes of a neural network (NN), to find the best training functions in classifying circular and organizational visualizations from our observers' pupillary responses.

2 Eye Tracking User Study

This experiment was conducted as part of an ARC Linkage project with Encompass Co. and aimed at collecting eye gaze data from observers when they interpret information visualization graphs. The two visualizations examined were the two most similar displays Encompass software commonly uses.

2.1 Experimental Procedure

The user study was conducted with two different types of graph visualizations. Here we considered a within subjects design, so we had to use different data inside the visualizations. We used circular and organizational visualizations, for which six very similar questions were asked for each as shown in Table 1. The Circular and Organizational visualizations are illustrated in Figures 1 and 2.

Table 1: The questions the observers were asked

No.	Question
Q1	Where does <a person> live?
Q2	Who lives at <a particular address>?
Q3	Who lives at <a different address>?
Q4	How many directors does <a particular company have>?
Q5	Which state do most of the directors of <a particular company> live in?
Q6	What connects <one company> to <another company>?

The two visualizations are apparently similar in appearance, but follow dissimilar principles [8]. The circular one highlights the structure of associations with a complex cluster of linked bodies, and the organizational one highlights each single body of major interest, for example the National Australia Bank or Sydney 2001 Olympics.

The duration of each question was 45 seconds. The visualizations and each question were displayed in an ordered way (a Latin square), so that it appeared different to each observer.

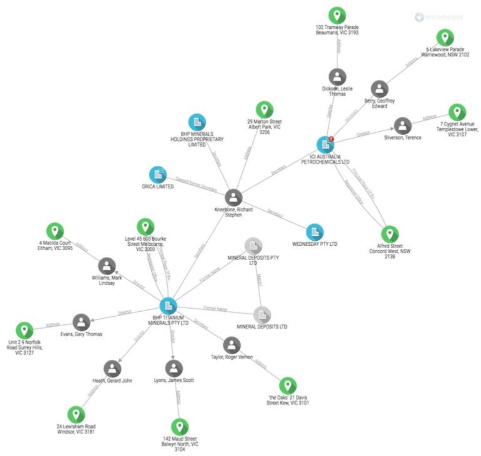

Figure 1: Circular Visualization.

Observers get a maximum of 45 seconds to answer each question, without a break between the two visualizations, and our software shows time remaining, and automatically move to the next question. All other properties are kept constant throughout the experiment such as the colours used and symbols, font size, graphical elements, and lighting environment. The questions are presented at the top left corner with an empty box to the right for typing the answer. The observers can see the visualizations while answering and if they finish early they can move to the next question by pressing the 'Next' button. Each experiment took about 12 minutes overall to be completed, including introduction, calibration, the two visualizations, and a closing discussion. A straightforward process was followed for the experiment without any other interaction between observer and experimenter.

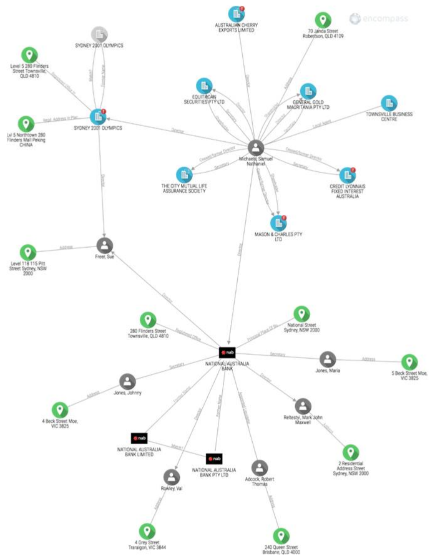

Figure 2: Organizational Visualization

2.2 Observers

Twenty-four students from the Australian National University (ANU) (11 male, 13 female) participated as observers in this experiment, with mean age of 22.1±5.5 (Mean ± SD) years. They had normal or corrected to normal vision. All participants signed an informed consent form prior to their participation. Prior to performing this experiment, ethics approval was received from the Australian National University's Human Research Ethics Committee. All observers participated voluntarily, among them 14 were undergraduate students who gained course credits for choosing to participate in experiments (chosen from a wise selection). No other incentive was involved in participating in this experiment.

2.3 Methodology

Pupillary responses of each observer were recorded using The Eye Tribe (https://theeyetribe.com/) remote eye-tracker system with a sampling rate of 60 Hz. A 15.6" Dell laptop and a computer mouse were peripherals for interaction between observer and the laptop running the web-based visualizations. Observer sat in a comfortable position in front of the laptop screen, and the distance between the eye tracker and observer was adjusted for best calibration results. In order to achieve the best results, observers were asked to restrict their body movement as much as possible.

Due to the nature of human natural physical movements and other effects, the recorded pupillary response is affected by noise like small signal fluctuations, eye blinking etc. Eye blinking points are zero in pupillary responses, so cubic spline interpolation and 10-point Hann moving window average are employed to reconstruct and smooth the pupil data respectively [9]. This procedure is applied to the left and right eyes' pupillary responses separately. Finally, maximum value normalization is applied to keep the signals in the range between 0 and 1.

Seven statistical temporal features, including amplitude (mean, maximum, minimum, range), standard deviation, means of the absolute values of the first and second derivatives of the processed signals, were extracted from pupillary responses. These features are very easy to compute and also cover the typical range, gradient and variation of the signals [10]. In this way, there were 84 extracted features (6 questions x 7 features x 2 pupils (left and right)) for an observer while watching either the circular or the organizational visualization.

In general, it is difficult to know which learning algorithm will be best suited for a given problem and this depends on different factors such as the number of the data points in the training set, the number of weights and biases on the network etc.

3 Results and Discussion

In this analysis, we used leave-one-observer-out process where the classifier was trained using twenty-three observers' data and tested by using the left-out observer's data. The average test performance over all observers is shown in Table 2. Here, 10 hidden nodes were considered during training of the network.

Table 2: Classification accuracies and elapsed time for different learning algorithms; Acc. = Accuracy, SD = Standard Deviation

Algorithms	Functions	Avg. Acc. (%)	SD	Time (Sec)
Gradient Descent	TRAINGD	56.3	21.7	21.8
	TRAINGDM	54.9	25.8	23.1
	TRAINGDA	80.2	22.4	22.3
	TRAINGDX	76.7	21.4	22.5
	TRAINRP	77.1	18.2	22.4
Conjugate Gradient	TRAINSCG	87.8	19.2	11.1
	TRAINCGP	79.9	18.2	19.2
	TRAINCGF	72.9	20.5	20.0
	TRAINCGB	70.5	22.9	19.0
Quasi-Newton	TRAINOSS	81.6	17.4	42.1
	TRAINBFG	74.0	22.6	09.1
	TRAINLM	**95.8**	**09.5**	**04.4**

It can be seen from Table 2 that the Levenberg-Marquardt (LM) algorithm shows the highest classification performance and lowest CPU elapsed time compared to the others. In this case, we find that the LM algorithm is more suitable in terms of classification accuracy and elapsed time.

There is evidence in the literature that this algorithm is beneficial when very accurate training is required, and contains up to a few hundred weights [11]. It has been necessary to focus on empirical analysis to find the best number of hidden nodes, as we expected and found that the number of hidden nodes affects classification accuracies [12]. In this case, we checked the effect of number of hidden nodes on the classification performance using LM. The variations of average classification accuracies with the various numbers of hidden nodes are shown in Figure 3 where error bars indicate standard deviations. We can see in Figure 3 that the average classification accuracy is improved with the increasing number of hidden nodes until 'enough' hidden nodes are used.

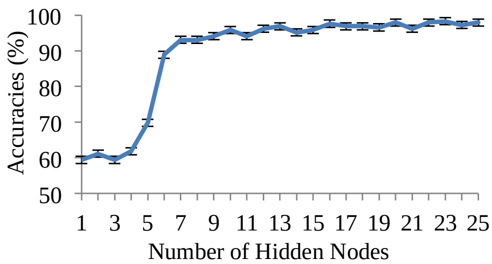

Figure 3: Variation of accuracies depending on the no. of hidden nodes of the neural network.

At this point, we wished to compare the observers' verbal responses with the classifier's performances. We trained a neural network using the Levenberg-Marquardt algorithm with 10 hidden nodes. The comparative analysis for the circular visualization is shown in Figure 4.

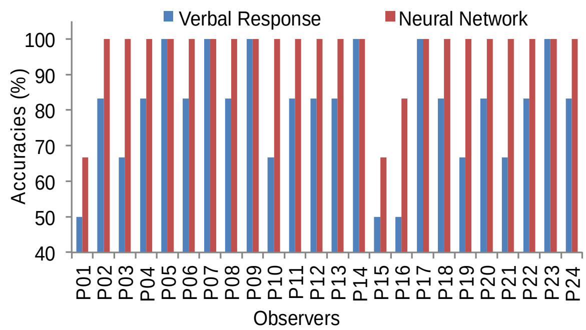

Figure 4: Performance of observers and classifier in case of Circular Visualization

We can observe from Figure 4 that the classifier shows higher accuracies in most of the observers compared to their own verbal responses (question answers are 'verbal' for short). We can see that *P01*, *P15*, and *P16* show lower accuracies compared to others, being 50% in case of verbal responses, and 66.7%, 66.7%, and 83.3% for pupillary responses, respectively. We can also see that, if observers were verbally 100% accurate, then the neural network is also 100% accurate for those observers. In an extended analysis, we found verbal responses and classifier's accuracies were correlated with a rate of 0.7035. On the average, observers were verbally 80.6% and physiologically 96.5% correct. The same procedure was applied to compare the performance with respect to the organizational visualisation as illustrated in Figure 5.

We can see from Figure 5 that the lowest verbal accuracies are found for *P01, P04, P05, P12, P13, P15, P19, P22,* and *P24,* being 66.7%, whereas the classifier's accuracies are 83.3%, 83.3%, 83.3%, 100%, 100%, 83.3%, 100%, 66.7%, and 83.3% for these identified observers, respectively. We can also see that when an observer is 100% correct, then the classifier's output is also 100% for that observer. In an extended analysis, we found that the correlation for the performance of observers and classifier is 0.6072. On the average, observers were verbally 81.3% and physiologically 95.1% correct.

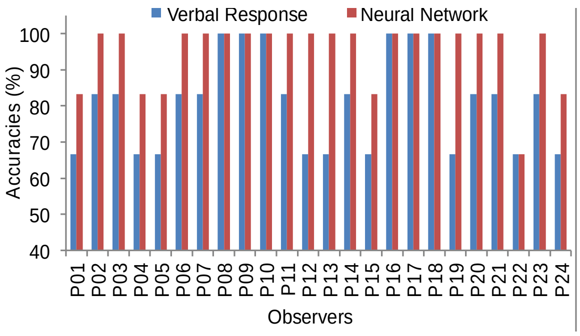

Figure 5: Performance of observers and classifier in case of Organizational Visualization.

We also adopted participants' responses as a target class and applied our neural network to classify them. It is worth mentioning that nine observers' verbal accuracy rate is higher for circular than organizational, eight participants' verbal accuracy rate is higher for organizational than circular where other seven observers are equally correct for both visualizations. As we adopted the classifier according to the participants' verbal response rate and applied our leave-one-observer-out approach, we get 84.7% and 83.3% classification accuracy rate for circular and organizational respectively. If we exclude the seven observers who are equally correct for both visualizations, then accuracy rate is improved to 96.1% and 95.1% for circular and organizational respectively which are very close to the our original results (96.5% and 95.1% for circular and organizational visualizations respectively). Thus, by considering all aspects, we can say that observers' physiological signals are more consistent while finding answers to questions from circular visualization compared to organizational visualization, and so our classifiers are more accurate.

4 Conclusion

In this study, observers are presented with six similar questions on two visualizations (circular and organizational). Observers' pupillary responses are recorded with remote eye tracking technology. Then, a neural network model is developed to classify these two visualizations from observers' pupillary responses and found that the Levenberg-Marquardt is the best algorithm in this case where seven or more hidden nodes are required to get higher accuracy. We also found that circular is better than organizational where only three observers did not answer questions correctly for the circular visualizations (*P01*, *P15*, and *P16*), whereas six observers did not answer correctly for the organizational visualization (*P01, P04, P05, P15, P22*, and *P24*). Two observers (*P01* and *P15*) showed relatively bad performances compared to others. A previous study has shown that the organizational visualization is more complex than the circular visualization [8], which is also found from our classifier performance. That is, observers are physiologically more accurate in the case of circular than organizational visualization, but vice versa in the case of verbal responses. Overall, it was found that observers are verbally 80.6% and 81.3% correct where they are physiologically 96.5% and 95.1% correct, for circular and organizational visualizations respectively. Thus, we can predict understanding of visualizations more accurately via unconscious pupillary responses compared to conscious verbal responses. This system will be useful to design user-adaptive visualizations considering users' physiological responses in addition to their verbal responses, which clearly cannot be assumed correct.

REFERENCES

[1]. M Tory and T Moller (2004). Human factors in visualization research. IEEE Transaction on Visualization and Computer Graphics, 10(1), 72-84.

[2]. A.A Lopes, R Pinho, F.V Paulovich, and R Minghim (2007). Visual text mining using association rules. Computers & Graphics, 31(3), 316-326.

[3]. P Pirolli, S.K Card, and M.M.V.D Wege (2003). The effects of information scent on visual search in the hyperbolic tree browser. ACM Transactions on Computer-Human Interaction, 10(1), 20-53.

[4]. A.V Moere, and H Purchase (2011). On the role of design in information visualization. Information Visualization, 10(4), 356-371.

[5]. K Sedig, S Rowhani, J Morey, H Liang (2016). Application of Information Visualization Techniques to the Design of a Mathematical Mindtool: A Usability Study. Information Visualization, 2(3), 142-156.

[6]. M Smiciklas, 2012. The Power of Infographics: Using Pictures to Communicate and Connect With Your Audiences. Que, Indianapolis, USA.

[7]. H Bruce, A Jocelyn, and M Tanya, 2016, Optimizing Visual Layout for Training and Learning Technologies, https://ro.uow.edu.au/acis2016/papers/1/12/.

[8]. M.Z Hossain, T Gedeon, S Caldwell, L Copeland, R Jones, and C Chow, 2018, Investigating differences in two visualisations from observer's fixations and saccades, https://dl.acm.org/citation.cfm?id=3167933

[9]. M.Z Hossain, T Gedeon, R Sankaranarayana, D Apthorp, A Dawel, 2016, Pupillary responses of Asian observers in discriminating real from fake smiles: A preliminary study, https://bit.ly/2Js3mrX.

[10]. R.W Picard, E Vyzas, and J Healey (2001). Toward machine emotional intelligence: analysis of affective physiological state. IEEE Transactions on Pattern Analysis and Machine Intelligence, 23(10), 1175-1191.

[11]. H.T. Gavin, 2017, The Levenberg–Marquardt method for nonlinear least squares curve-fitting problems. Duke University, https://bit.ly/2F5uTkD.

[12]. W Zou, Y. Li, and A. Tang, 2009. Effects of the number of hidden nodes used in a structured-based neural network on the reliability of image classification. Neural Computing and Applications, 18(3), 249-260.

[13]. D Mulligan, 2010. i2 Analyst Notebook: A Visual Step by Step Guide of Introductory and Advanced Charting Techniques, https://bit.ly/2yA7IsQ

[14]. Cambridge Intelligence, KeyLines Layouts, 2018. https://bit.ly/ 2CJYME2