A Field Dependence-Independence Perspective on Eye Gaze Behavior within Affective Activities

Christos Fidas[1], Marios Belk[2,4], Christodoulos Constantinides[4], Argyris Constantinides[3,4], Andreas Pitsillides[4]

[1]Department of Electrical and Computer Engineering
University of Patras, Patras, Greece
fidas@upatras.gr
[2]Cognitive UX GmbH, Heidelberg, Germany
belk@cognitiveux.de
[3]Cognitive UX LTD, Nicosia, Cyprus
argyris@cognitiveux.com
[4]Department of Computer Science, University of Cyprus, Nicosia, Cyprus
cconst04@cs.ucy.ac.cy, andreas.pitsillides@cs.ucy.ac.cy

Abstract. Evidence suggests that human cognitive differences affect users' visual behavior within various tasks and activities. However, a human cognitive processing perspective on the interplay between visual and affective aspects remains up-to-date understudied. In this paper, we aim to investigate this relationship by adopting an accredited cognitive style framework (Field Dependence-Independence – FD-I) and provide empirical evidence on main interaction effects between human cognition and emotional processing towards eye gaze behavior. For doing so, we designed and implemented an eye tracking study (*n*=22) in which participants were initially classified according to their FD-I cognitive processing characteristics, and were further exposed to a series of images, which triggered specific emotional valence. Analysis of results yield that affective images had a different effect on FD and FI users in terms of visual information exploration time and comprehension, which was reflected on eye gaze metrics. Findings highlight a hidden and rather unexplored effect between human cognition and emotions towards eye gaze behavior, which could lead to a more holistic and comprehensive approach in affective computing.

Keywords: Individual Differences, Cognitive Processing Styles, Global and Local Processing, Human Emotions, Eye Tracking, User Study.

1 Introduction

Human emotions are essential in human experience, influencing perception, as well as logical and rational thinking in cognitive processing [1, 2]. Researchers and practitioners in the area of Affective Computing have studied the importance of human emotions in human-computer interaction [2], interactive systems' design [3], and adaptive user interfaces [4] that personalize their behavior and responses according to the users' current emotional states.

An important challenge in such research endeavors relates to the non-intrusive and implicit elicitation of users' emotional states. For this purpose, several research works have investigated whether users' visual behavior can be a predictor of their emotional states during memory- and decision-making tasks [5-11]. Lemos et al. [5] introduced an automated method for measuring (un)pleasant emotions and the level of excitement to a stimulus before they are cognitively perceived and interpreted by the human mind by analyzing the gaze properties of a person (eye gaze, blinks, pupil change). Schmid et al. [6] studied how mood states affect information processing during facial emotion recognition by analyzing the users' global and local visual information processing styles. Sears et al. [7] analyzed and investigated visual attention to emotional images in previously depressed individuals. Stanley et al. [8] studied the interplay among cultural differences, visual behavior and emotions utilizing emotional images. Charoenpit and Ohkura [9] explored emotions within eye gaze-driven e-learning systems aiming to detect curiosity, interest and boredom based on low-level eye gaze metrics such as fixations, fixations' duration and pupil diameter. Zheng et al. [10] and Zhen-Fen et al. [11] proposed novel emotion recognition methods by combining electroencephalography (EEG) signals with users' pupillary responses and scan-paths, respectively.

Research Motivation. Visual processing is performed through two distinct information streams of the human brain; the *global* and the *local information streams* [12, 13]. *Global processing* enables individuals to process the visual field holistically to give meaning to an object or a scene [12], while *local processing* enables individuals to process the details out of a whole [13]. Humans process information cognitively through both the global and local information streams by filtering relevant information within a visual scene [14]. Although individuals integrate both streams during cognitive processing, research indicates that *individual differences in cognitive processing styles* exist, which highlight the individuals' preferred mode of visually processing information either *globally (holistically)* or *locally (analytically)* [15]. In addition, prior research suggests that when individuals recognize emotions, they process stimuli globally rather than locally [6, 19, 20], *i.e.*, the stimulus is primarily processed as a whole and not in a fragmented fashion.

Unlike existing works [6, 20], which investigated how mood states affect global and local information processing, in this work, we go a step further by investigating whether there is a systematic effect of human cognitive differences towards eye gaze behavior within affective activities, framed by an accredited human cognition theory (*i.e.*, Witkin's Field Dependence-Independence [15]). Considering such an interplay could lead to a more comprehensive approach in affective computing. To the best of our knowledge, such an interplay has not been investigated so far.

The remainder of the paper is structured as follows. First, we present the underlying theory on human cognitive differences and human emotions. We further present the method of study, analysis of results, discussion on the main findings, limitations, and future steps.

2 Background Theory

Among several cognitive processing style theories, this work focuses on the *Field Dependence-Independence cognitive style theory* [15], which is an accredited and widely applied theory [16, 17, 18, 38] that holistically represents how individuals perceive their context and surrounding field in which a task needs to be accomplished. It highlights human cognitive differences into *Field Dependent (FD) (or Wholists)* and *Field Independent (FI) (or Analysts)*. Evidence has shown that FD and FI individuals have differences in visual perceptiveness [15], visual working memory capacity [16], visual search abilities [17], in the way they organize and process information of their surrounding visual field [18]. In particular, FD individuals view the perceptual field as a whole and they are not attentive to detail, while FI individuals view discrete information presented by their visual field as a collection of parts.

Furthermore, affective experiences such as feelings and mood can be defined as combinations of two basic dimensions: *valence* and *arousal* [21, 22]. *Valence* describes the attractiveness or aversiveness of stimuli along a continuum (negative-neutral-positive), while *arousal* relates to the perceived intensity of an event ranging from very calming to highly exciting or agitating [22, 23]. Research has shown that cognitive processing characteristics of individuals affect the way individuals control their emotions, *e.g.*, individuals with enhanced cognitive processing abilities, control their emotions more naturally than those with more limited cognitive abilities [28]. Several works also indicate that emotional arousal affects cognition such as the long-term memory [23, 24].

3 Method of Study

3.1 Research Questions

RQ$_1$. Are there significant differences in users' visual exploration time between FD *vs.* FI users for images that trigger different emotional valence?

RQ$_2$. Are there significant differences in users' eye gaze transition entropy between FD *vs.* FI users for images that trigger different emotional valence?

3.2 Apparatus

The research instruments utilized in the study include: *i)* the Group Embedded Figures Test (GEFT) paper-and-pencil test for classifying the participants into FD and FI groups; *ii)* a wearable eye tracking device; *iii)* a wearable electroencephalography (EEG) device for emotion recognition; and *iv)* a conventional desktop computer.

Human Cognitive Style Elicitation. Users' field dependence-independence was measured through the GEFT [29], which is a widely applied and validated paper-and-pencil test [16, 17, 18]. The test measures the user's ability to find common geometric shapes in a larger design. The GEFT consists of 25 items. In each item, a simple geometric figure is hidden within a complex pattern, and participants are required to identify the simple figure by drawing it with a pencil over the complex figure. Based on a widely applied cut-off score [16, 32], participants that solve less than 12 items are FD, while participants that solve 12 items and above are FI.

Equipment. An All-in-One HP computer with a 24" monitor was used (1920x1080 pixels). To capture eye movements, we used Pupil Labs Core eye tracker [30], which captures data at 200Hz. The eye tracker was calibrated individually using a 5-point calibration procedure, and was positioned at an upwards angle roughly 30cm below eye level and 65cm away. To capture human emotions, we used Emotiv Epoc+ [31], a 14-channel electroencephalography, which detects emotional states, such as excitement, engagement, relaxation, interest, stress, focus.

3.3 Sampling and Procedure

A total of 22 individuals (9 females) participated, aged between 20-32 years old (*m*=24, *std*=3.1). Participants were split into two groups based on the GEFT classification (12 FD, 10 FI).

Aiming to control the effect of image content on creating emotions (positive/negative/neutral), we used an open-source data set of emotionally annotated images (*i.e.*, Open Affective Standardized Image Set (OASIS) [33]). Furthermore, we asked users to think aloud by describing: *i)* the image (to trigger visual exploration and search); *ii)* the feelings the image triggers; and *iii)* the reason the particular feeling was triggered. Throughout the session, participants wore the eye tacking and EEG device for measuring the eye gaze behavior and emotional states respectively (both devices were calibrated at the beginning of each user session). Such a procedure allowed us to triangulate results from subjective feedback (users' expressing their feelings on emotional valence – positive/neutral/negative), and objective measures derived from Emotiv in which the detected emotional states were post-processed and assigned to an emotional valence group (valence – positive/neutral/negative).

Considering that users' cognitive styles and image complexity affect users' visual behavior [32, 34], we chose images of similar complexity between and within images belonging to the three groups. For doing so, we assessed the equivalence of the two image sets by calculating the image complexity using entropy estimators [35, 36]. Following a within-subjects study design, all participants were presented with a series of the same images that were randomly assigned to the participants.

We adopted the University's human research protocol that takes into consideration users' privacy, confidentiality and anonymity. All participants performed the task in a quiet lab room with only the researchers present.

3.4 Image Set

To investigate the research questions, we intentionally chose three image sets: *i) Positive Images:* images that trigger positive valence; *ii) Negative Images:* images that trigger negative valence; and *iii) Neutral Images:* images that trigger neutral valence. For this purpose, we have used the Open Affective Standardized Image Set (OASIS) [33], an open-access online stimulus set containing 900 colored images depicting a broad spectrum of themes, including humans, animals, objects, and scenes along with normative ratings on emotional valence and arousal. **Figure 1** illustrates the study images, and **Table 1** summarizes the emotional valence and arousal triggered for each image according to OASIS, and the estimated image complexity.

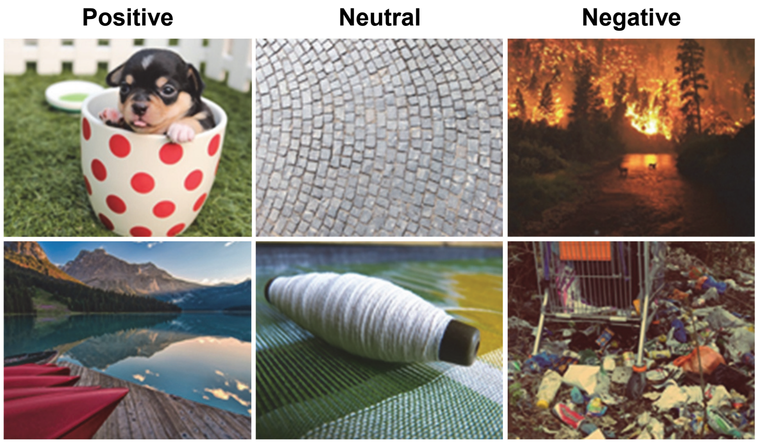

Fig. 1. The set of images used in the study and the corresponding emotions that are triggered.

Table 1. Emotional valence/arousal according to OASIS, and image complexity for each image used in the study.

Image	Emotional Valence	Emotional Arousal	Complexity in bits
Dog	6.49 (positive)	5.02	7.45
Lake	6.41 (positive)	4.10	7.87
Sidewalk	4.29 (neutral)	2.23	7.34
Yarn	4.20 (neutral)	1.98	7.65
Fire	1.75 (negative)	5.31	7.54
Garbage	1.63 (negative)	4.78	7.23

3.5 Data Metrics

For measuring *visual exploration time*, we measured the time that started as soon as the image was illustrated to the users, until the users expressed loudly their feelings. For capturing the users' *eye gaze behavior*, we used raw fixation metrics (count and duration) using Pupil Labs' built-in methods. We have further extended the software to analyze Areas of Interests (AOI), and gaze transition entropy (based on [37]) to assess whether a user is performing a careful view of AOIs *vs.* more randomness and more frequent switching between AOIs.

4 Analysis of Main Effects

There were no significant outliers in the data based on boxplot inspection, and residuals were normally distributed. Data are mean ± standard deviation.

4.1 Visual Exploration Time Differences between FD *vs.* FI Users for Images that trigger Different Emotional Valence (*RQ₁*)

A two-way mixed analysis of variance was conducted to examine interaction effects among cognitive style (FD/FI) and image type (positive/neutral/negative) on the time to explore the image (**Figure 2**). There was no statistically significant interaction between cognitive style and image type on the time to explore the image, $F(2, 60)=.955$, $p=.391$, *partial* $\eta^2=.031$, while there was an effect of image type on the time to explore the image, $F(2, 60)=7.572$, $p=.001$, *partial* $\eta^2=.202$. Furthermore, there was a significant difference between the time to explore images between FD and FI users, $F(1, 30)=5.551$ $p=.025$, *partial* $\eta^2=.156$. Descriptive statistics reveal that affective images (positive and negative) triggered higher exploration times compared to neutral images across cognitive style groups. FD users explored for a longer time all three image types compared to FI users.

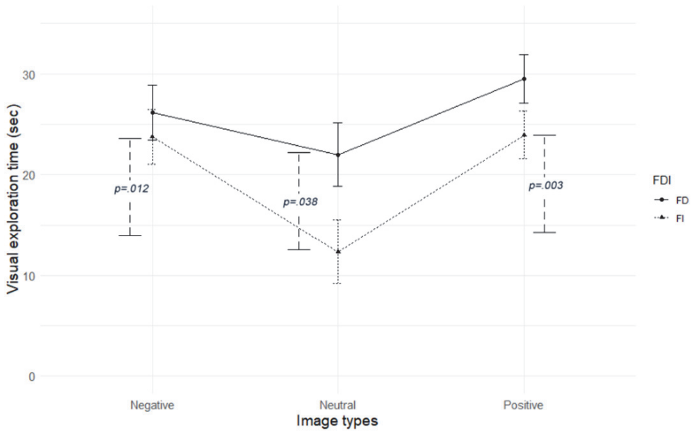

Fig. 2. Mean visual exploration times (sec) among user groups.

Main Finding related to *RQ₁*. FD users spent more time viewing affective (positive and negative valence) images than FI users (**Figure 2**). Such an observation can be explained by the fact that FD users have a more trained global information processing stream compared to FI users [15, 16, 17, 18, 25, 38], and consequently interpreted semantically faster the provided image, which is reflected on their visual behavior. In addition, FD users inherently follow a more holistic and exploratory approach as opposed to FI users that primarily focus their attention on areas of interests, hence, FD users spent more time exploring the image compared to FI users.

4.2 Eye Gaze Behavior Differences between FD *vs.* FI Users for Images that trigger Different Emotional Valence (*RQ2*)

A two-way mixed analysis of variance was conducted to examine interaction effects among cognitive style (FD/FI) and image type (positive/neutral/negative) on fixation count and fixation duration to assess visual behavior (**Table 2**). There was no statistically significant interaction between cognitive style and image type on fixation count, $F(2, 60)=1.397$, $p=.255$, *partial* $\eta^2=.044$, nor on fixation duration, $F(2, 60)=.626$, $p=.538$, *partial* $\eta^2=.02$. Nevertheless, descriptive statistics indicate that for FIs, fixation count and fixation duration was amplified in affective images (positive and negative) compared to neutral images, whereas for FDs, fixation count and duration was similar across image types with a tendency of higher scores in affective images.

Table 2. Mean fixation count and duration among user groups.

Image	Fixation Count		Fixation Duration (sec)	
	FD (std.)	FI (std.)	FD (std.)	FI (std.)
Positive	24.18 (3.27)	20.68 (2.67)	2.91 (1.50)	2.09 (1.96)
Negative	22.75 (3.99)	23.62 (1.70)	2.59 (1.40)	2.25 (1.55)
Neutral	20.37 (4.31)	12.25 (4.49)	2.71 (2.10)	1.34 (1.67)

To further assess visual search behavior, we compared transition entropy among user groups (**Figure 3**). Based on Krejtz et al. [37], lower values of transition entropy H_t indicate more careful/focused viewing of Areas of Interests (AOI), while greater H_t values indicate more randomness and more frequent switching between AOIs. There was a statistically significant interaction between cognitive style and image type on transition entropy, $F(2, 60)=3.417$, $p=.039$, *partial* $\eta^2=.102$. An analysis of simple main effects shows a statistically significant difference in mean transition entropy among FD and FI users in the positive images, $F(1, 30)=5.023$, $p=.033$, as well as for the negative images, $F(1, 30)=4.522$, $p=.042$, but not for the neutral images ($p>.05$).

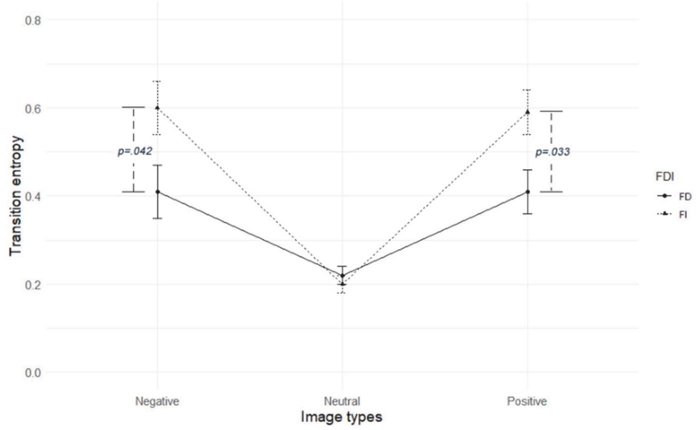

Fig. 3. Mean transition entropy among user groups.

Main Finding related to *RQ_2*. FI users switched on a more global visual information processing approach in affective images (positive and negative) compared to neutral images in which they followed their inherent local stream of visual information processing (**Table 2**) in which they inherited a more analytic approach. While FI users inherently follow an analytic visual exploration approach by using the local information processing stream, in affective images they tend to use their global processing stream as revealed through the analysis of eye gaze metrics. In addition, FD and FI users' visual strategies were amplified significantly on affective images as reflected on eye gaze behavior in terms of eye gaze transition entropy (**Figure 3**), which paves the way for considering cognitive style as an important factor in affective computing.

5 Conclusion

In this paper we investigated whether users' inherent and preferred ways of information processing are affected during affective activities, triggered by visual information. For doing so, we conducted an in-lab eye tracking study in which we classified users into Field Dependent (FD) and Field Independent (FI) based on an accredited human cognition theory and cognitive elicitation test, and further exposed them to a series of images, which are known to trigger specific emotional valence.

Beyond our expectations analysis of results point towards a swapping effect in the case of affective images between FD and FI users. In particular, FI users followed a rather holistic approach as opposed to the typical analytical approach, while FD users where not significantly affected. Such an endeavor could pave the path for a new research paradigm for taking into consideration a more holistic and comprehensive approach in interactive system design. Specifically, our work contributes on considering human cognitive differences as an important prediction factor in order to estimate through real-time eye gaze analysis techniques the emotional valence of end-users during affective activities. For example, application domains such as gaming, mixed reality, etc. could deploy intelligent human cognitive and affective elicitation mechanisms (*e.g.*, [26, 27, 39]) by considering the main effects reported in this study. Nonetheless, more research is needed to manifest this approach with more empirical data.

Limitations relate to the small sample size and controlled lab setting. In addition, a specific image set was used, which however was necessary to better control the experiment. Future work entails embracing images with varying complexity, and images that trigger higher/lower levels of arousal to increase external validity of this work.

Acknowledgements. This research has been partially supported by the EU Horizon 2020 Grant 826278 "Securing Medical Data in Smart Patient-Centric Healthcare Systems" (Serums), the Research and Innovation Foundation (Project DiversePass: COMPLEMENTARY/0916/0182), and the European project TRUSTID - Intelligent and Continuous Online Student Identity Management for Improving Security and Trust in European Higher Education Institutions (Grant Agreement No: 2020-1-EL01-KA226-HE-094869), which is funded by the European Commission within the Erasmus+ 2020 Programme and the Greek State Scholarships Foundation I.K.Y.

References

1. Picard, R. (1997). Affective Computing. M.I.T. Press, Cambridge, MA
2. Picard, R. (1999). Affective Computing for HCI. Conference on Human-Computer Interaction: Ergonomics and User Interfaces-Volume I. Lawrence Erlbaum Associates
3. Norman, D. (2003). Emotional Design: Why we Love (or Hate) Everyday Things. Basic Books
4. Natasha, J., Conati, C., Harley, J., Azevedo, R. (2014). Predicting Affect from Gaze Data during Interaction with an Intelligent Tutoring System. Intelligent Tutoring Systems
5. Lemos, J., Sadeghnia, G., Ólafsdóttir, Í., Jensen, O. (2008). Measuring Emotions using Eye Tracking
6. Schmid, P., Mast, M., Mast, F., Lobmaier, J. (2011). How Mood States Affect Information Processing during Facial Emotion Recognition: An Eye Tracking Study. Swiss Journal of Psychology, 70, 223-231
7. Sears, C., Kristin N., Ference, J. Thomas, C. (2011). Attention to Emotional Images in Previously Depressed Individuals: An Eye-Tracking Study. Cognitive Therapy and Re-search, 35, 517-528
8. Stanley, J., Zhang, X., Fung, H., Isaacowitz, D. (2013). Cultural Differences in Gaze and Emotion Recognition: Americans Contrast more than Chinese. Emotion, 13(1), 36-46
9. Charoenpit, S., Ohkura, M. (2014). Exploring Emotion in an E-Learning System using Eye Tracking. Symposium on Computational Intelligence in Healthcare and E-Health, 141-147
10. Zheng, W., Dong, B., Lu, B. (2104). Multimodal Emotion Recognition using EEG and Eye Tracking Data. Conference on Engineering in Medicine and Biology Society, 5040-5043
11. Zhen-Fen, S., Chang, Z., Wei-Long, Z., Bao-Liang, L. (2017). Attention Evaluation with Eye Tracking Glasses for EEG-based Emotion Recognition. Neural Engineering, 86-89
12. Hübner, R. Volberg, G. (2005). The Integration of Object Levels and Their Content: A Theory of Global/Local Processing and Related Hemispheric Differences. Journal of Experimental Psychology: Human Perception and Performance, 31(3), 520-541
13. Oliva, A. (1997). Coarse Blobs or Fine Edges? Evidence That Information Diagnosticity Changes the Perception of Complex Visual Stimuli. Cognitive Psychology, 34 (1), 72-107
14. Davidoff, J., Fonteneau, E., Fagot, J. (2008). Local and Global Processing: Observations from a Remote Culture. Cognition, 108 (3), 702-709
15. Witkin, H., Moore, C., Goodenough, D., Cox, P. (1975). Field-Dependent and Field-Independent Cognitive Styles and their Educational Implications. Research Bulletin, 1-64
16. Hong, J., Hwang, M., Tam, K., Lai, Y., Liu, L. (2012). Effects of Cognitive Style on Digital Jigsaw Puzzle Performance: A GridWare Analysis. Computers in Human Behavior, 28(3), 920-928
17. Rittschof, K. (2010). Field Dependence-Independence as Visuospatial and Executive Functioning in Working Nemory: Implications for Instructional Systems Design and Research. Educational Technology Research and Development, 58(1), 99-114
18. Belk, M., Fidas, C., Germanakos, P., Samaras, G. (2017). The interplay between humans, technology and user authentication: a cognitive processing perspective. Computers in Human Behavior, 76, Elsevier, 184-200
19. Calder, A., Young, A., Keane, J., Dean, M. (2000). Configural Information in Facial Expression Perception. Exp. Psych.: Human Perception & Performance, 26, 527–551.
20. Prkachin, G. (2003). The Effect of Orientation on Detection and Identification of Facial Expressions of Emotion. British Journal of Psychology, 94, 45-62

21. Watson, D., Wiese, D., Vaidya, J., Tellegen, A. (1999). The Two General Activation Systems of Affect: Structural Findings, Evolutionary Considerations, and Psychobiological Evidence. Journal of Personality and Social Psychology, 76(5), 820
22. Kensinger, E., Schacter, D. (2006). Processing Emotional Pictures and Words: Effect of Valence and Arousal. Cogn. Affect. Behav. Neurosci., 6, 110-126
23. Costanzi, M., Cianfanelli, B., Saraulli, D., Lasaponara, S., Doricchi, F., Cestari, V., Rossi-Arnaud, C. (2019). The Effect of Emotional Valence and Arousal on Visuo-Spatial Working Memory: Incidental Emotional Learning and Memory for Object-Location. Frontiers in Psychology
24. Cahill, L., McGaugh, J. (1995). A Novel Demonstration of Enhanced Memory Associated with Emotional Arousal. Consciousness and Cognition, 4(4), 410-421
25. Belk, M., Fidas, C., Germanakos, P., Samaras, G. (2015). Do human cognitive differences in information processing affect preference and performance of CAPTCHA?. International Journal of Human-Computer Studies, 84, Elsevier, 1-18
26. Constantinides, A., Belk, M., Fidas, C., Pitsillides, A. (2020). An eye gaze-driven metric for estimating the strength of graphical passwords based on image hotspots. ACM IUI 2020, ACM Press, 33-37
27. Constantinides, A., Belk, M., Fidas, C., Pitsillides, A. (2019). On the Accuracy of Eye Gaze-driven Classifiers for Predicting Image Content Familiarity in Graphical Passwords. ACM UMAP ’19, ACM Press, 201–205.
28. Schmeichel, B., Demaree, H. (2010). Working Memory Capacity and Spontaneous Emotion Regulation: High Capacity Predicts Self-enhancement in Response to Negative Feedback. Emotion, 10, 739-744
29. Oltman, P., Raskin, E., Witkin, H. (1971). A Manual for the Embedded Figures Test. Consulting Psychologists Press, Palo Alto CA, USA
30. Kassner, M., Patera, W., Bulling, A. (2014). Pupil: An Open Source Platform for Pervasive Eye Tracking and Mobile Gaze-based Interaction. ACM UbiComp 2014 Adjunct, ACM Press, 1151-1160
31. Emotiv Epoc+ (2021). Available online: https://www.emotiv.com/epoc
32. Katsini, C., Fidas, C., Raptis, G., Belk, M., Samaras, G., Avouris, N. (2018). Influences of Human Cognition and Visual Behavior on Password Security during Picture Password Composition. ACM Human Factors in Computing Systems (CHI 2018), ACM Press, p. 87
33. Kurdi, B., Lozano, S., Banaji, M. (2017). Introducing the Open Affective Standardized Image Set (OASIS). Behavior Research Methods, 49, 457-470
34. Raptis, G., Katsini, C., Belk, M., Fidas, C., Samaras, G., Avouris, N. (2017). Using Eye Gaze Data and Visual Activities to Infer Human Cognitive Styles: Method and Feasibility Studies. Conference on User Modeling, Adaptation and Personalization, 164-173
35. Cardaci, M., Di Gesù, V., Petrou, M., Tabacchi, M. (2009). A fuzzy Approach to the Evaluation of Image Complexity. Fuzzy Sets and Systems, 160(10), 1474-1484
36. SciKit Image (2021). Available online at https://scikit-image.org/docs/dev/api/skimage.measure.html#skimage.measure.shannon_entropy
37. Krejtz, K., Duchowski, A., Szmidt, T., Krejtz, I., Perilli, F.G., Pires, A., Vilaro, A., Villalobos, N. (2015). Gaze Transition Entropy. ACM Trans. Appl. Percept. 13, 1, article 4
38. Fidas, C., Belk, M., Hadjidemetriou, G., Pitsillides, A. (2019). Influences of mixed reality and human cognition on picture passwords: An eye tracking study. IFIP TC13 Human-Computer Interaction (INTERACT 2019), Springer-Verlag, 304-313
39. Costi, A., Belk, M., Fidas, C., Constantinides, A., Pitsillides, A. (2020). CogniKit: An extensible tool for human cognitive modeling based on eye gaze analysis. ACM Intelligent User Interfaces (IUI Companion 2020), ACM Press, 130-131