Studying the Effect of Human Cognition on Text and Image Recognition CAPTCHA Mechanisms

Marios Belk[1], Panagiotis Germanakos[1], Christos Fidas[1], George Spanoudis[2], and George Samaras[1]

[1] Department of Computer Science, University of Cyprus, CY-1678 Nicosia, Cyprus
{belk,pgerman,christos.fidas,cssamara}@cs.ucy.ac.cy
[2] Department of Psychology, University of Cyprus, CY-1678 Nicosia, Cyprus
spanoud@ucy.ac.cy

Abstract. This paper investigates the effect of individual differences in human cognition on user performance in CAPTCHA tasks. In particular, a three-month ecological valid user study was conducted with a total of 107 participants who interacted with a text-recognition and an image-recognition CAPTCHA mechanism. The study included a series of psychometric tests for eliciting users' speed of processing, controlled attention and working memory capacity, with the aim to examine the effect of these cognitive processes on the efficiency and effectiveness of user interactions in CAPTCHA tasks. Preliminary results provide interesting insights for the design and deployment of adaptive CAPTCHA mechanisms based on individual differences in cognitive processing since it has been initially shown that specific cognitive processing abilities of individuals could be a determinant factor on the personalization of CAPTCHA mechanisms.

Keywords: Individual Differences, Cognitive Processing Abilities, CAPTCHA, Efficiency, Effectiveness, User Study.

1 Introduction

A CAPTCHA (Completely Automated Public Turing test to tell Computers and Humans Apart) [1] is a computer program widely used today for protecting Web applications against automated software agents whose purpose is to degrade the quality of a provided service, whether due to misuse or resource expenditure. A typical example of a text-based CAPTCHA challenge (Figure 1) verifies that the entity interacting with a remote service is a human being, and not a machine, by requiring from a legitimate user to type letters or digits based on a distorted image that appears on the screen. Such a challenge is based on the assumption that a distorted text-based image can be easily recognized by the human brain but present significant difficulty for optical character or image recognition systems.

Research on CAPTCHA mechanisms has received significant attention lately with the aim to improve their usability and at the same time prevent adversarial attacks by malicious software. Researchers promote various CAPTCHA designs based on

L. Marinos and I. Askoxylakis (Eds.): HAS/HCII 2013, LNCS 8030, pp. 71–79, 2013.

text- and speech-recognition challenges, and image puzzle problems [2, 3, 4]. Nevertheless, a variety of studies have been reported that underpin the necessity for improving the usability of CAPTCHA mechanisms [5, 6, 7, 8]. Results from a recent study, which investigated users' perceptions towards CAPTCHA challenges, claim that current implementations do not provide an acceptable trade off solution with regards to CAPTCHA usability [5]. Another large-scale study which evaluated CAPTCHAs on the Internet's biggest Web-sites revealed that CAPTCHAs are difficult for humans to solve [6].

Fig. 1. Example of a text-recognition CAPTCHA

Taking into consideration that human computer interactions with regard to CAPTCHA mechanisms are in principal cognitive tasks that embrace to recognize and process information, we suggest that these interactions should be analyzed in more detail under the light of cognitive theories. Theories of individual differences in human cognition aim to describe and explain how and why individuals differ in cognitive abilities [9, 10]. In this respect, various researchers attempted to explain the functioning of the human mind in terms of more basic processes, such as speed of processing, controlled attention and working memory capacity [11]. *Speed of processing* refers to the maximum speed at which a given mental act may be efficiently executed [12, 13]. *Controlled attention* refers to cognitive processes that can identify and concentrate on goal-relevant information and inhibit attention to irrelevant stimuli [12, 14]. *Working memory capacity* is defined as the maximum amount of information that the mind can efficiently activate during information processing as an empirical model of cognitive functions used for temporarily storing and manipulating information [15, 16].

To this end, given that the aforementioned cognitive factors have a main effect on mental tasks, such as information processing, comprehension, learning, and problem solving [9, 15], we suggest that such characteristics should be utilized as part of an adaptive interactive system specialized in personalizing CAPTCHA tasks to the cognitive processing abilities of each user. In this respect we further describe the results and findings of a user study that aimed to investigate whether there is a main effect of users' cognitive processing abilities, targeting on speed of processing, controlled attention and working memory capacity, on the efficiency and effectiveness of different types of CAPTCHA mechanisms. In particular, a text- and an image-recognition CAPTCHA mechanism were deployed in the frame of an ecological valid experimental design, to investigate the effect of cognitive processing abilities of individuals towards efficiency and effectiveness with regard to CAPTCHA tasks.

The paper is structured as follows: next we describe the context of the empirical study and its methodology. Thereafter, we analyze and discuss the findings of the study. Finally, we summarize our findings and outline the implications of the reported research.

2 Method of Study

2.1 Procedure

A Web-based environment was developed within the frame of various university courses which was used by the students throughout the semester as an online blog for posting comments related to the course, as well as for accessing the courses' material (i.e., course slides, homework, etc.) and for viewing their grades. The participants were required to solve CAPTCHA challenges throughout the semester primarily before posting comments on the online blog. In particular, participants were randomly provided with different variations of CAPTCHA mechanisms (i.e., text-recognition or image-recognition). For example, in case a user solved a text-recognition CAPTCHA at time 0, the system would provide the same user to solve an image-recognition CAPTCHA at time 1 in the future with the aim to engage the whole sample with different types of CAPTCHA.

The text-recognition mechanism was developed using available open-source software that produced distorted images of random characters[1]. Furthermore, we have utilized Microsoft ASIRRA (Animal Species Image Recognition for Restricting Access) [3] as the image-recognition CAPTCHA mechanism that produced pictures and asked the participants to select the appropriate pictures belonging to a specific group (i.e., select pictures that illustrate cats among dogs). Figure 2 and Figure 3 respectively illustrate the text- and image-recognition CAPTCHA mechanisms utilized in the study.

Fig. 2. Text-recognition CAPTCHA used in the study

Both client-side and server-side scripts were developed to monitor the users' behavior during interaction with the CAPTCHA mechanism. In particular, the total time (efficiency) and the total number of attempts (effectiveness) required for successfully solving the CAPTCHA challenge were monitored on the client-side utilizing a browser-based logging facility that started recording time as soon the CAPTCHA challenge was presented to the users until they successfully completed the CAPTCHA task. For user identification, the Web-site further utilized the participants' username since the course's Web-site required user authentication for accessing the course's material.

---

[1] Securimage v. 3.0, `http://www.phpcaptcha.org`.

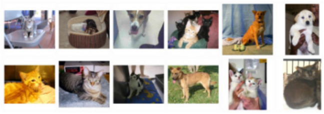

Fig. 3. Image-recognition CAPTCHA used in the study

Controlled laboratory sessions were also conducted throughout the period of the study to elicit the users' cognitive factors (speed of processing, controlled attention and working memory capacity) through a series of psychometric tests [9, 15, 16]. With the aim to apply the psychometric tests in a scientific right manner, we conducted several sessions with a maximum of 5 participants by following the protocol suggested by the inventors of the psychometric tests. The psychometric tests utilized in the study are described next.

Users' Speed of Processing Elicitation Test. A Stroop-like task was devised to measure simple choice reaction time to address speed of processing. Participants were instructed to read a number of words denoting a color written in the same or different ink color (e.g., the word "red" written in red ink color). A total of 18 words were illustrated to the participants illustrating the words "red", "green" or "blue" either written in red, green or blue ink color. The participants were instructed to press the R key of the keyboard for the word "red", the G key for the word "green" and the B key for the word "blue". The reaction times between 18 stimuli and responses onset were recorded and their mean and median were automatically calculated.

Users' Controlled Attention Elicitation Test. Similar to the speed of processing elicitation test, a Stroop-like task was devised, but instead of denoting the word itself, participants were asked to recognize the ink color of words denoting a color different than the ink (e.g., the word "green" written in blue ink). A total of 18 words were illustrated to the participants illustrating the words "red", "green" or "blue" either written in red, green or blue ink color. The participants were instructed to press the R key of the keyboard for the word written in red ink color, the G key for the word written in green ink color and the B key for the word written in blue ink color. The reaction times between 18 stimuli and responses onset were recorded and their mean and median were automatically calculated.

Users' Working Memory Capacity Elicitation Test. A visual test addressed storage capacity in short-term memory [15, 16]. In particular, the psychometric test illustrated a geometric figure on the screen and the participant was required to memorize the figure. Thereafter, the figure disappeared and 5 similar figures were illustrated on the screen, numbered from 1 to 5. The participant was required to provide the number (utilizing the keyboard) of the corresponding figure that was the same as the initial figure. The test consisted of 21 figures (seven levels of three trials each). As the participant correctly identified the figures of each trial, the test provided more complex figures as the levels increased indicating an enhanced working memory capacity.

2.2 Participants

The study was conducted between September and November 2012 with a total of 107 participants (52 male, 55 female, age 17-26, mean 22). Participants were undergraduate students of Computer Science, Electrical Engineering, Psychology and Social Science departments.

2.3 Analysis of Results

Overarching aim of the study was to investigate whether there is a significant difference with regard to time (efficiency) and total number of attempts (effectiveness) needed to solve a text- and image-recognition CAPTCHA mechanism among users with different cognitive processing abilities. For our analysis, we separated participants into different groups based on their cognitive processing abilities (limited, intermediate, enhanced) of each cognitive factor (speed of processing, controlled attention, working memory capacity).

CAPTCHA Solving Efficiency. A series of three by two way factorial analyses of variance (ANOVA) were conducted aiming to examine main effects of users' cognitive processing differences (i.e., limited, intermediate, enhanced) and CAPTCHA type (i.e., text- and image-recognition) on the time needed to accomplish the CAPTCHA task. Figure 4 illustrates the means of performance per cognitive factor group in regard with the speed of processing (SP), controlled attention (CA) and working memory capacity (WMC) dimension, and CAPTCHA type.

Results revealed that there is a main effect of the speed of processing and controlled attention dimensions on the time needed to solve a CAPTCHA challenge (SP: $F(1,424)=3.819$, $p=0.023$; CA: $F(1,424)=28.889$, $p=0.029$). On the other hand, no safe conclusions can be drawn at this point in time whether there is a main effect of working memory capacity of users on the time needed to solve a CAPTCHA challenge (WMC: $F(1,424)=1.172$, $p=0.311$) since users across all three groups did not perform significantly different in both types of CAPTCHA challenges. Accordingly, these findings suggest that speed of processing and controlled attention abilities primarily affect the users' interactions with CAPTCHA challenges, whereas in the case of working memory, results indicate that different capacities of working memory may affect performance, however not significantly. Such a result might be based on the fact that enhanced speed of processing and controlled attention is needed to efficiently focus a person's attention on the distorted characters among the added noise of current text-recognition CAPTCHAs, as well as the recognition of particular objects in image-recognition CAPTCHAs.

A further comparison between CAPTCHA types (text- vs. image-recognition) for each cognitive processing dimension revealed that users with enhanced cognitive processing abilities performed significantly faster in text-recognition CAPTCHAs than image-recognition CAPTCHAs (SP: $F(1,151)=12.155$, $p=0.001$; CA: $F(1,160)=13.751$, $p<0.001$; WMC: $F(1,142)=13.375$, $p<0.001$). On the other hand, users with intermediate and limited speed of processing and controlled attention, no significant differences were observed between solving efficiency in text- and image-recognition CAPTCHAs. In this respect, from a user-adaptation point of view, an adaptive CAPTCHA mechanism could recommend users with intermediate and limited speed of processing and controlled attention abilities an image-recognition CAPTCHA as an alternative security solution to the currently dominant text-recognition CAPTCHAs, with the aim to provide an improved user experience. In the same context, in the case of users with enhanced speed of processing and controlled attention, the adaptive CAPTCHA mechanism could recommend a text-recognition CAPTCHA given that users with enhanced cognitive processing abilities performed significantly faster in text- than in image-recognition CAPTCHAs.

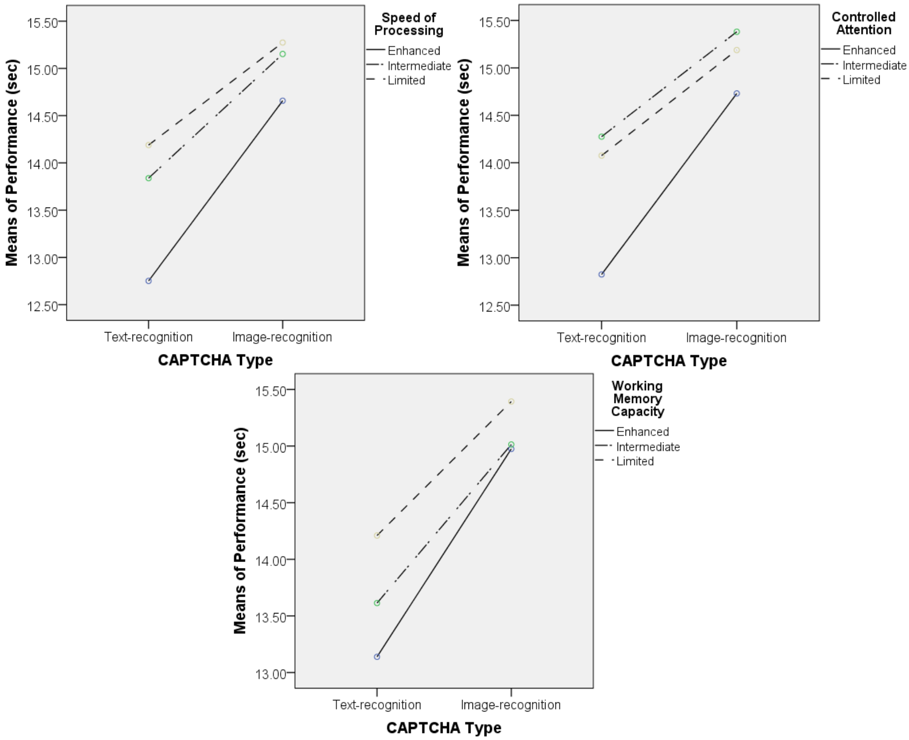

Fig. 4. Means of Performance for all three Cognitive Processing User Groups

Regarding the working memory capacity dimension, it is yet not clear whether it could be used as a personalization factor since results did not reveal significant differences between user groups for both types of CAPTCHA. In this respect, an adaptive CAPTCHA mechanism that would decide on a CAPTCHA type according to this user characteristic would not be able to clearly distinguish performance differences among groups of users.

CAPTCHA Solving Effectiveness. For each user session the total number of attempts made for successfully solving the CAPTCHA challenge was recorded. Table 1 and Table 2 respectively summarize the means of attempts per cognitive processing group (i.e., SP, CA, WMC groups) for each CAPTCHA type (i.e., text- and image-recognition). Shapiro-Wilk tests revealed that these distributions do not follow the normal distribution. In the case of text-recognition CAPTCHAs, on average, users with limited CA and limited WMC needed more attempts to solve the CAPTCHA challenges than the other two groups (intermediate and enhanced groups). The Kruskal-Wallis test revealed that the differences between controlled attention users was statistically significant ($H(2)=9.167$, $p=0.001$), as well as in the case of working memory capacity users ($H(2)=6.464$, $p=0.039$). In the case of the speed of processing user group, no significant differences have been observed between number of attempts of each user group, as the Kruskal-Wallis test revealed ($H(2)=3.744$, $p=0.154$), suggesting that this cognitive dimension might not significantly affect the effectiveness of CAPTCHA.

Table 1. Means of Attempts per User Group for Text-recognition CAPTCHA

	Speed of Processing		**Controlled Attention**		**Working Memory Capacity**	
	Mean	**Std. Dev.**	**Mean**	**Std. Dev.**	**Mean**	**Std. Dev.**
Enhanced	2	1.37	1.71	1.07	1.82	1.06
Intermediate	1.71	1.1	1.38	0.81	1.51	0.98
Limited	1.6	1.05	2.1	1.42	2.21	1.66

In the case of image-recognition CAPTCHAs, no significant differences in solving effectiveness have been observed between the user groups since the majority of image-recognition CAPTCHAs were solved at first attempt across all user groups, indicating that cognitive processing abilities might not primarily affect user effectiveness in image-recognition CAPTCHA tasks.

Table 2. Means of Attempts per User Group for Image-recognition CAPTCHA

	Speed of Processing		**Controlled Attention**		**Working Memory Capacity**	
	Mean	**Std. Dev.**	**Mean**	**Std. Dev.**	**Mean**	**Std. Dev.**
Enhanced	1.5	0.59	1.17	0.57	1.26	0.45
Intermediate	1.13	0.35	1.21	0.41	1.35	0.56
Limited	1.17	0.38	1.46	0.5	1.3	0.48

To this end, initial findings indicate that differences in controlled attention and working memory capacity might affect the effectiveness of text-recognition CAPTCHA challenges since users of the intermediate and enhanced user groups needed less attempts than the ones of the limited user groups. On the other hand, no safe conclusions can be drawn whether there is a main effect of users' cognitive processing abilities on the solving effectiveness of image-recognition CAPTCHAs since no clear differences have been recorded between solving attempts among user groups.

3 Conclusions

This paper reported the results of a three-month ecological valid user study that entailed credible psychometric-based tests for eliciting users' cognitive characteristics and two variations of CAPTCHA mechanisms (text- and image-recognition), with the aim to investigate whether individuals with different cognitive processing abilities perform differently in terms of efficiency and effectiveness in CAPTCHA tasks.

Initial results demonstrate a main effect of cognitive processing abilities primarily in solving efficiency of text-recognition CAPTCHA mechanisms. In particular, results revealed that users with enhanced controlled attention and speed of processing performed significantly faster than users with limited processing abilities in text-recognition CAPTCHAs. A comparison between text- and image-recognition CAPTCHAs revealed that users with enhanced cognitive processing abilities performed significantly faster in the text-recognition CAPTCHAs, however, users with intermediate and limited speed of processing and controlled attention did not significantly perform differently in text- than image-recognition CAPTCHAs. Given that no significant differences were observed in this case, users with limited and intermediate cognitive processing abilities could benefit with an image-recognition CAPTCHA than text-recognition. Regarding effectiveness (total number of attempts), initial findings indicate that differences in controlled attention and working memory capacity might affect the effectiveness of text-recognition CAPTCHA challenges since users with limited cognitive processing abilities needed significantly more attempts than the other two user groups. On the other hand, speed of processing has not affected the effectiveness of solving CAPTCHA since the differences among user groups were not significant. Furthermore, in the case of image-recognition CAPTCHAs, given that the majority of user interactions solved the challenge at first attempt, results suggest that cognitive processing abilities do not strongly affect image-recognition in these particular CAPTCHA tasks.

The limitations of the reported study are related to the fact that participants were only university students with an age between 17 to 26 years. In this respect, further studies need to be conducted with a greater sample of varying profiles and ages in order to reach to more concrete conclusions about the effect of individuals' cognitive processing abilities on their performance in CAPTCHA challenges. On the other hand, there has been an effort to increase ecological and internal validity of the research since the CAPTCHA tasks were integrated in a real Web-based system and the participants were involved at their own physical environments without the intervention of any experimental equipment or person.

The majority of CAPTCHAs utilized today on the Internet are primarily based on text-recognition challenges [17]. The results of this study suggest enhancing current CAPTCHA mechanisms to embrace both text- and image-recognition CAPTCHA challenges. Such an endeavor would have many positive implications on the usability and user experience of security-related interactions since adapting CAPTCHA challenges based on individual differences in cognitive processing could improve the effectiveness and efficiency of such tasks, minimize the users' added cognitive loads and learning efforts, as well as minimize erroneous interactions in CAPTCHA tasks.

Acknowledgements. The work is co-funded by the PersonaWeb project under the Cyprus Research Promotion Foundation (ΤΠΕ/ΠΛΗΡΟ/0311(ΒΙΕ)/10), and the EU projects Co-LIVING (60-61700-98-009) and SocialRobot (285870).

References

1. von Ahn, L., Blum, M., Langford, J.: Telling Humans and Computers Apart Automatically. Communications of the ACM 47, 56–60 (2004)
2. von Ahn, L., Maurer, B., McMillen, C., Abraham, D., Blum, M.: reCAPTCHA: Human-Based Character Recognition via Web Security Measures. Science 321(5895), 1465–1468 (2008)
3. Elson, J., Douceur, J., Howell, J., Saul, J.: Asirra: A CAPTCHA that Exploits Interest-Aligned Manual Image Categorization. In: 14th ACM Conference on Computer and Communications Security, pp. 366–374. ACM Press, New York (2007)
4. Vikram, S., Fan, Y., Gu, G.: SEMAGE: a New Image-based Two-factor CAPTCHA. In: 27th ACM Conference on Computer Security Applications, pp. 237–246. ACM Press, New York (2011)
5. Fidas, C., Voyiatzis, A., Avouris, N.: On the Necessity of User-friendly CAPTCHA. In: 29th ACM Conference on Human Factors in Computing Systems, pp. 2623–2626. ACM Press, New York (2011)
6. Bursztein, E., Bethard, S., Fabry, C., Mitchell, J.C., Jurafsky, D.: How Good are Humans at Solving CAPTCHAs? A Large Scale Evaluation. In: IEEE International Symposium on Security and Privacy, pp. 399–413. IEEE Press, Washington (2010)
7. Yan, J., El Ahmad, A.S.: Usability of CAPTCHAs or Usability Issues in CAPTCHA Design. In: 4th ACM Symposium on Usable Privacy and Security, pp. 44–52. ACM Press, New York (2008)
8. Hernandez-Castro, C., Ribagorda, A.: Pitfalls in CAPTCHA Design and Implementation: The Math CAPTCHA, A Case Study. J. Computers and Security 29, 141–157 (2010)
9. Demetriou, A., Spanoudis, G., Shayer, S., Mouyi, A., Kazi, S., Platsidou, M.: Cycles in Speed-Working Memory-G Relations: Towards a Developmental-Differential Theory of the Mind. J. Intelligence 41, 34–50 (2013)
10. Hunt, E.B.: Human Intelligence. Cambridge University Press, New York (2011)
11. Demetriou, A., Spanoudis, G., Mouyi, A.: Educating the Developing Mind: Towards an Overarching Paradigm. J. Educational Psychology Review 23(4), 601–663 (2011)
12. MacLeod, C.M.: Half a Century of Research on the Stroop Effect: An Integrative review. J. Psychological Bulletin 109, 163–203 (1991)
13. Posner, M.I., Raicle, M.E.: Images of Mind. Scientific American Library, New York (1997)
14. Stroop, J.R.: Studies of Interference in Serial Verbal Reactions. J. Experimental Psychology 18, 643–662 (1935)
15. Baddeley, A.: Working Memory: Theories, Models and Controversies. Annual Review of Psychology 63, 1–29 (2012)
16. Baddeley, A.: Working Memory. J. Science 255(5044), 556–559 (1992)
17. Bursztein, E., Martin, M., Mitchell, J.: Text-based CAPTCHA Strengths and Weaknesses. In: 18th ACM International Conference on Computer and Communications Security, pp. 125–138. ACM Press, New York (2011)