Towards the Personalization of CAPTCHA Mechanisms based on Individual Differences in Cognitive Processing

Marios Belk[1], Panagiotis Germanakos[1,2], Christos Fidas[1], Andreas Holzinger[3], George Samaras[1]

[1]Department of Computer Science, University of Cyprus, CY-1678 Nicosia, Cyprus
[2]SAP AG, Dietmar-Hopp-Allee 16, 69190 Walldorf, Germany
{belk,pgerman,christos.fidas,cssamara}@cs.ucy.ac.cy,
[3]Institute for Medical Informatics, Statistics & Documentation, Research Unit HCI4MED
Medical University Graz, A-8036 Graz, Austria
andreas.holzinger@medunigraz.at

Abstract. This paper studies the effect of individual differences on user performance related to text-recognition CAPTCHA challenges. In particular, a text-recognition CAPTCHA mechanism was deployed in a three-month user study to investigate the effect of individuals' different cognitive processing abilities, targeting on speed of processing, controlled attention and working memory capacity toward efficiency and effectiveness with regard to different levels of complexity in text-recognition CAPTCHA tasks. A total of 107 users interacted with CAPTCHA challenges between September and November 2012 indicating that the usability of CAPTCHA mechanisms may be supported by personalization techniques based on individual differences in cognitive processing.

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

1 Introduction

One of the most important security concerns on the World Wide Web today is to protect Web-based systems and services against automated software agents whose purpose is to degrade the quality of a provided service. Examples include among others the automated creation of fake email accounts that are used later on for spam, generation of massive scale advertising, manipulation of online voting systems, access of private information, generation of hyperlinks in forums to improve their Web-sites' search engine ranking, dictionary attacks of passwords, etc.

A "Completely Automated Public Turing test to tell Computers and Humans Apart" (CAPTCHA) [1, 2] is a widely used security mechanism to protect Web applications, interfaces, and services from such malicious software by verifying that the entity interacting with a system is actually a human being, and not a machine. A typical example of a CAPTCHA mechanism requires from a legitimate user to type letters or digits based on a distorted image that appears on the screen (Figure 1). 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 an optical character or image recognition system.

CAPTCHA challenges over the World Wide Web are performed primarily with the use of text-recognition CAPTCHA [3, 4, 5]. The reCAPTCHA project [4], which is currently the most popular and widely used CAPTCHA online, estimates that over 200 million reCAPTCHAs are completed daily, and it takes an average of 10 seconds to complete one. In addition, major Web service providers such as Google, Facebook, Microsoft and many others utilize text-recognition CAPTCHA to protect their premises against automated attacks [3].

Fig. 1. Example of a text-recognition reCAPTCHA

Current text-based CAPTCHA implementations suffer from an important drawback; making the characters of the CAPTCHA hard to be recognized by computer systems, also increases the difficulty for humans, and thus decreases usability of interaction. The problem is further strengthened by the improvement of current character recognition systems that are more capable of breaking CAPTCHA mechanisms [6, 7], and as a consequence, the characters' distortion and complexity is increased making it even more difficult to be recognized by humans. Various studies have been reported that underpin the necessity for increasing usability of current CAPTCHA implementations. A study in [8] raised the usability issues of CAPTCHA and proposed a framework for evaluating various designs. A recent study which investigated users' perceptions toward CAPTCHA challenges underpins the necessity for user friendly CAPTCHA challenges as current implementations do not provide an acceptable trade off solution with regard to usability [9].

Within this realm, given that individuals share different characteristics, needs and preferences, supporting usability of CAPTCHA systems with adaptation and personalization technologies [10] may improve the system's usability and user experience by providing users with adaptive and personalized CAPTCHA challenges according to their unique characteristics. Given that current text-based CAPTCHA implementations require from individuals to recognize specific characters among irrelevant, noisy information, and process this information on a cognitive level, we suggest that individual differences in cognitive processing should be taken into consideration in the design of current text-recognition CAPTCHA mechanisms. In this respect, the purpose of this paper is to investigate the effect of specific individual characteristics of users targeting on cognitive processes (i.e., speed of processing, controlled attention and working memory capacity), toward efficiency and effectiveness of different variations of text-recognition CAPTCHA challenges in terms of complexity (i.e., low, medium and high level of complexity illustrating respectively, 5, 7 or 9 characters with increased character distortion and noise).

2 Individual Differences in Adaptive Interactive Systems

Overarching aim is to drive this research towards the design and the deployment of adaptive and personalized CAPTCHA mechanisms that will assist users to accomplish efficiently and effectively usable CAPTCHA tasks. In this respect we provide an overview of adaptive interactive systems to elicit how these could be of value for designing adaptive CAPTCHA mechanisms.

An adaptive interactive system [10] is any interactive system which is capable to automatically or semi-automatically adapt its information architecture and functionality as a response on implicit or explicit gathered data which are related with the users themselves, their interaction with the system or the context of use in which interaction takes place. The utter goal is to increase the functionality of the system and improve the users' experience by providing personalized and bootstrapped functionalities.

One distinctive feature of an adaptive interactive system is its user model. The user model is a representation of static and dynamic information about an individual, and it represents an essential entity for an adaptive interactive system aiming to provide adaptation effects (i.e., the same system can look different to users with different user models). For example, an information retrieval system can adaptively select and prioritize the most relevant items to the user's goals and/or interests. A security task in a commercial Web system can present the content adaptively to the user's level of knowledge towards security terms (e.g., provide novice users with personalized security information awareness by using additional explanations).

Adaptive interactive systems build and maintain a data model throughout computer human interaction which entails information considered essential in order to adapt content and functionalities to the unique characteristics of a user. According to the nature of information that is being modeled, we distinguish models that represent information about the user (user's knowledge, interests, goals, background and personality traits) and about the user's context of use (user's location, platform, physical environment) [10].

A considerable amount of research efforts have been undertaken focusing on modeling and utilizing personality traits (e.g., cognitive factors) for personalization in adaptive interactive systems. Several works [11, 12, 13] have distinguished users based on their cognitive characteristics, and provided different adaptation effects accordingly. A study in [11] distinguished imager and verbal users, and wholist and analyst users based on Riding's Cognitive Style Analysis [14], as well as based on the users' cognitive processing abilities. Each user was provided with adaptive presentation of content (graphical or verbal), different navigation organization (stressing analytic or holistic navigation approaches) and different amount of user control and content based on the users' cognitive processing abilities. In a similar approach in [12] users were distinguished in field-dependent and field-independent based on Witkin et al. [15] and provided different navigation organization, amount of user control, and navigation support tools for these groups. Results of these studies indicate that cognitive characteristics have significant impact in the adaptation and personalization process of Web environments by increasing usability and user satisfaction during navigation as well as learning performance.

Such findings suggest that individual differences in human cognition are important to take into account in the personalization process of an adaptive interactive system. Accordingly, modeling and adapting CAPTCHA mechanism based on users' cognitive factors could improve CAPTCHA solving efficiency and effectiveness, and minimize users' cognitive loads and erroneous interactions by providing different levels of complexity according to the cognitive characteristics of each individual.

In this context, a number of theories in human cognition exist that aim to describe and explain [16, 17, 18] the functioning of the human mind in terms of more basic processes, such as *speed of processing* which indicates the time needed by the human mind to record and give meaning to information [19, 20], *controlled attention* which refers to cognitive processes that can identify and concentrate on goal-relevant information and inhibit attention to irrelevant stimuli [19, 21], and *working memory capacity* which refers to a cognitive system for temporary storage of information and information manipulation [22, 23]. Various research works argue that these cognitive processes have an effect on comprehension, learning and problem solving [24-29]. They are mainly used in mental tasks, such as arithmetic tasks; remembering a number in a multiplication problem and adding that number later on, or creating a new password and using that password later for authentication, or recognizing the distorted text of a CAPTCHA mechanism.

To this end, given that the aforementioned cognitive factors have a main effect in problem solving and other tasks (e.g., individuals with increased working memory capacity accomplish tasks more efficiently), we suggest that such characteristics should be utilized as part of an adaptive interactive system specialized in personalizing CAPTCHA-related tasks to the cognitive abilities of each user. Main aim of this paper is to investigate how individuals with differing cognitive processing abilities perform in various CAPTCHA challenges. Such an endeavor could support the development of an adaptation engine that would embrace cognitive characteristics as its core user model and accordingly adapt the level of complexity in each CAPTCHA solving task to improve usability of interactions.

3 Method of Study

3.1 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) and for viewing their grades. The participants were required to solve text-recognition CAPTCHA challenges with different levels of complexity throughout the semester, primarily before posting comments on the online blog. In particular, a text-recognition CAPTCHA challenge with low, medium or high level of complexity was provided to each user. For example, in case a user solved a CAPTCHA with medium level of complexity at time 0, the system would provide the same user to solve a CAPTCHA with different level of complexity (low or high complexity) at time 1 in the future with the aim to engage the whole sample with all available levels of CAPTCHA complexity.

With the aim to increase the internal validity of the study, the different levels of complexity were designed specifically to ensure that the relationship among the three different CAPTCHA types regarding their complexity level was linear. In particular, the criteria for developing the different levels of complexity were based on the number of characters presented, and the percentage of text distortion and noise illustrated in each CAPTCHA challenge. The low complex CAPTCHA entailed 5 characters and 40% text distortion and noise, while the medium and high complex CAPTCHAs entailed respectively 7 and 9 characters, and 50% and 60% text distortion and noise, as illustrated in Figure 3.

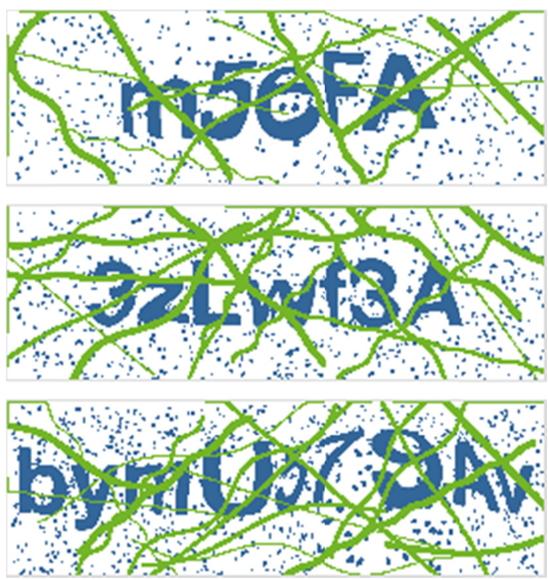

Fig. 2. Text-recognition CAPTCHA with Increasing Complexity

The text-recognition mechanism was developed using available open-source software that produced distorted images of random characters[1]. The text-recognition CAPTCHA mechanism also contained a refresh button that initialized the CAPTCHA with a new sequence of characters. Both client-side and server-side scripts were developed to monitor the users' behavior during interaction with the CAPTCHA mech-

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

anism. In particular, the total time (efficiency) and total number of attempts and refresh events (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 process. For user identification, the Web-site utilized the participants' username since the course's Web-site required user authentication for accessing the course's material.

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 [23, 26]. The psychometric tests utilized in the study are described next.

Users' Speed of Processing Elicitation Test. The test required participants to read a number of words designating a color written in the same or different ink color (e.g., the word "blue" written in blue ink color). Eighteen 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 eighteen stimuli and responses were recorded and their mean and median were automatically calculated (as suggested in [26]).

Users' Controlled Attention Elicitation Test. A similar test to the previous one was utilized, but instead of denoting the word itself, participants were required to recognize the ink color of words denoting a color different than the ink (e.g., the word "blue" written in green ink). Again, eighteen words were illustrated to the participants illustrating the words "red", "green" or "blue" either written in red, green or blue ink color, and the participants had to respond as quick as possible utilizing the keyboard. The reaction times between eighteen stimuli and responses were recorded and their mean and median were automatically calculated (as suggested in [26]).

Users' Working Memory Capacity Elicitation Test. Two tasks addressed storage capacity in short-term memory, the verbal and the visual test [23], whose results were combined to indicate a user's working memory capacity. The *visual test* illustrated a geometric figure on the screen and participants were required to memorize the figure. Thereafter, the figure disappeared and five similar figures were illustrated on the screen, numbered from one to five (Figure 4). Participants were required to provide the number of the figure illustrating the same shape as the initial figure through the keyboard. The test consisted of twenty one figures (seven levels of three trials each). As participants correctly identified the figures, the test provided more complex figures indicating an enhanced working memory capacity.

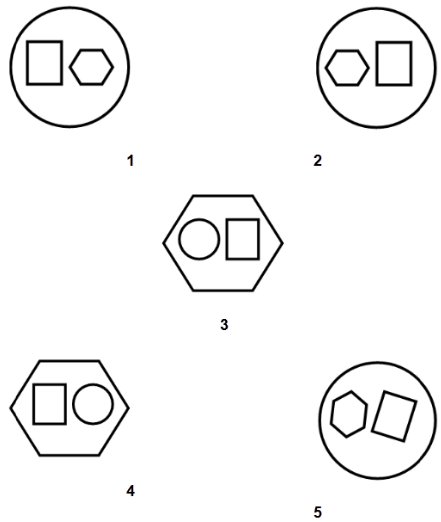

Fig. 3. Visual Working Memory Test

The *verbal test* showed a series of statements and required the participants to respond whether they are true or false. In addition, participants were required to remember the last word of each sentence and then write the last word of the sentence. The test included six levels of difficulty, e.g., in level three, participants were required to respond true/false to three successive sentences and had to remember and provide the last word of each sentence. For example, for the sentences "Knives are sharp", "The sun is shining", and "Fish have fur" the participant should respectively respond true, true and false, and then provide the word "sharp", "shining" and "fur" to the system (Figure 5). The level each participant reached indicated his/her working memory capacity.

Fig. 4. Verbal Working Memory Test

3.2 Participants

A total of 107 undergraduate students (52 male, 55 female, age 17-26, mean 22) participated in the study. A total of 1172 CAPTCHA sessions have been recorded during a three-month period.

3.3 Hypothesis

The following hypothesis was formulated for the purpose of our research:

***H_1*.** There is significant difference with regard to time (efficiency), total number of attempts and total number of refresh events (effectiveness) needed to solve a CAPTCHA mechanism among users with different cognitive processing abilities.

3.4 Analysis of Results

For the analysis, we separated participants into different categories based on their cognitive processing characteristics (limited, intermediate, enhanced) of each cognitive factor (speed of processing, controlled attention, working memory capacity), which are summarized in Table 1.

Table 1. User Groups based on Cognitive Processing Abilities

	Speed of Processing		**Controlled Attention**		**Working Memory Capacity**	
	F	**%**	**F**	**%**	**F**	**%**
Enhanced	59	55.1	33	30.8	46	43
Intermediate	19	17.8	18	16.8	27	25.2
Limited	29	27.1	56	52.3	34	31.8
Total	107	100	107	100	107	100

CAPTCHA Solving Efficiency. A series of three by three 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 complexity (i.e., low, medium, high) on the time needed to accomplish the CAPTCHA task. Figure 6 and Figure 7 respectively illustrate the means of performance per cognitive factor group in regard with the speed of processing (SP) and controlled attention (CA) dimension, and CAPTCHA complexity level. Table 2 and Table 3 respectively summarize post-hoc tests with CAPTCHA performance comparisons between each cognitive factor's user group (SP and CA).

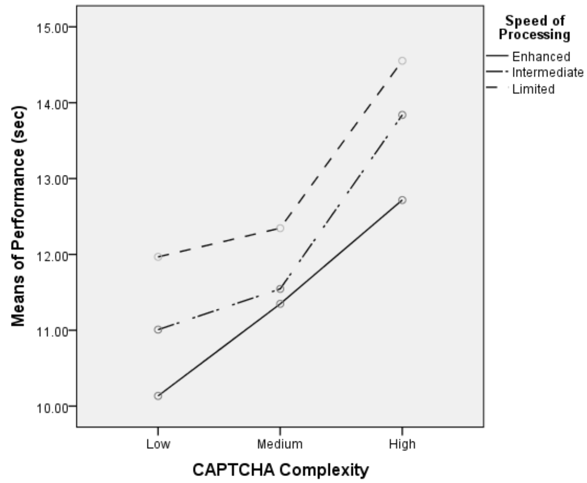

Fig. 5. Means of Performance for Speed of Processing User Groups

The results revealed that there is a significant main effect on the time needed to solve a CAPTCHA challenge with regard to both speed of processing (SP) and controlled attention (CA) factors, and CAPTCHA complexity level. Users with enhanced SP solved significantly faster all three complexity types of CAPTCHA compared to the limited user group (p<.001).

Table 2. Multiple Comparisons between the Speed of Processing User Group

		Mean Difference	Std. Error	Sig.
Enhanced	Intermediate	-0.60	0.38	0.117
	Limited	-1.07	0.27	0.000
Intermediate	Enhanced	0.60	0.38	0.117
	Limited	-0.46	0.36	0.202
Limited	Enhanced	1.07	0.27	0.000
	Intermediate	0.46	0.36	0.202

On the other hand, users with intermediate SP did not perform significantly different compared to the limited user group (p=.202) and the enhanced user group (p=.117). Based on Figure 6, given that the means of performances of enhanced SP users in the high complex CAPTCHA, and limited SP users in the medium complex CAPTCHA were not significantly different, an adaptive CAPTCHA mechanism embracing these user characteristics could provide a highly complex CAPTCHA mechanism to users with enhanced SP, and a medium complex CAPTCHA mechanism to users with limited SP, thus having increased security for enhanced SP users, but at similar levels of usability in terms of efficiency compared to limited SP users. Similarly, given that users with limited and intermediate SP did not perform significantly different in low and medium levels of CAPTCHA complexity, in order to increase security, a medium level of complexity could be used at a minimum cost to usability.

Regarding the controlled attention (CA) dimension, results similarly indicate that users with enhanced CA perform significantly faster compared to users with limited CA (p=.001), however with no significant differences compared to users with intermediate CA (p=.111). On the other hand, no significant differences were observed between the limited and intermediate user groups across all three types of CAPTCHA complexity level (p=0.444).

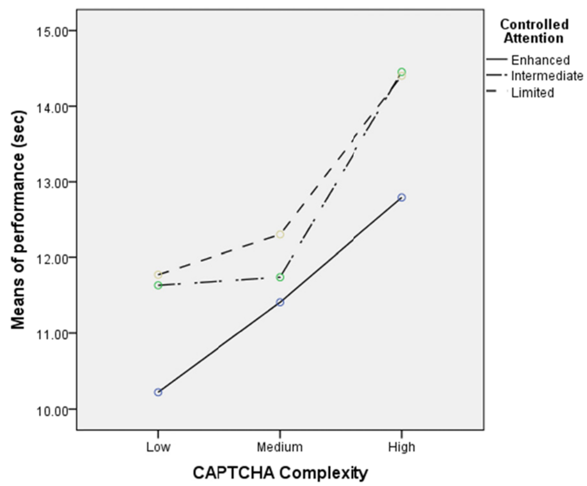

Fig. 6. Means of Performance for Controlled Attention User Groups

From a user-adaptation point of view, an adaptive CAPTCHA mechanism could recommend a CAPTCHA with medium complexity to users with limited and intermediate CA, given that these groups did not perform significantly different in the low and medium complex CAPTCHA. Regarding users with enhanced CA, a highly complex CAPTCHA for increased security could be provided since they did not perform significantly different in highly complex CAPTCHA compared to medium complex CAPTCHA interactions of the other two user groups, thus increasing security at minimum cost to usability for users with enhanced CA.

Table 3. Multiple Comparisons between the Controlled Attention User Group

		Mean Difference	**Std. Error**	**Sig.**
Enhanced	Intermediate	-0.63	0.39	0.111
	Limited	-0.93	0.27	0.001
Intermediate	Enhanced	0.63	0.39	0.111
	Limited	-0.29	0.39	0.444
Limited	Enhanced	0.93	0.27	0.001
	Intermediate	0.29	0.39	0.444

Regarding the working memory capacity (WMC) dimension, results also revealed a main effect of working memory capacity of users on the time to solve CAPTCHAs with different levels of complexity. Figure 8 illustrates the means of performance per WMC user group and CAPTCHA level of complexity. Table 4 summarizes the post-hoc tests with CAPTCHA performance comparisons between users with different working memory capacity.

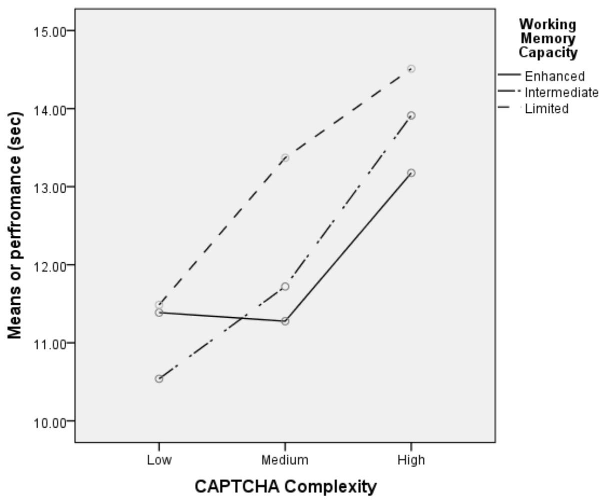

Fig. 7. Means of Performance for Working Memory Capacity User Groups

Accordingly, users with limited WMC performed significantly slower than users with intermediate WMS ($p=.002$) and enhanced WMC ($p<.001$). However, in the case of low level of CAPTCHA complexity we observe that users with enhanced WMC performed similarly as in the case of limited WMC users. A within analysis of the enhanced WMC group in this case revealed that the majority of users had limited and intermediate speed of processing which might have affected their performance. Such an observation is in line with previous theory suggesting that enhanced speed of information processing facilitates access to information that is sustained in the working memory system, and thus decreased speed of processing of users might have negatively affected their efficiency of information manipulation in the working memory system [23]. On the other hand, no significant differences between users with enhanced and intermediate WMC were observed.

To this end, as in the case of SP and CA factors, results suggest that individual differences in working memory capacity should be considered in CAPTCHA designs since a main effect of working memory capacity of users on time to solve CAPTCHA challenges has been observed. Results suggest that a CAPTCHA challenge with medium complexity could be provided to users with intermediate and enhanced WMC since no significant differences in performance were observed compared to the low complex (and less secure CAPTCHA), increasing thus security at a minimum cost to usability. In the case of limited WMC users, results suggest providing a less secure CAPTCHA to increase usability. However, regarding the intermediate and enhanced WMC groups, significant differences in their performance has been observed between the medium and high levels of CAPTCHA complexity, suggesting that a less complex (medium level), but more usable CAPTCHA challenge would significantly benefit the users.

Table 4. Multiple Comparisons between the Working Memory Capacity User Group

		Mean Difference	**Std. Error**	**Sig.**
Enhanced	Intermediate	-0.20	0.27	0.455
	Limited	-1.35	0.37	0.000
Intermediate	Enhanced	0.20	0.27	0.455
	Limited	-1.15	0.36	0.002
Limited	Enhanced	1.35	0.37	0.000
	Intermediate	1.15	0.36	0.002

CAPTCHA Solving Effectiveness based on Number of Attempts. The effectiveness of a solver measures their accuracy, i.e., the fraction of challenges they answered correctly [30]. For each user session the total number of attempts made for successfully solving the CAPTCHA challenge was recorded. Table 5 summarizes the means of attempts across all three levels of CAPTCHA complexity per cognitive processing group (i.e., SP, CA, WMC groups). Shapiro-Wilk tests revealed that these distributions do not follow the normal distribution. 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 5. Means of Attempts per User Group [35]

	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

Furthermore, the majority of CAPTCHA sessions (59.6%) of the sample were solved at the first try across all three cognitive groups. As expected, in the case of the low complexity CAPTCHA, 71.4% of the cases were solved at the first try, whereas in the case of medium and high complexity, the percentage decreased to 57.1% and 54.3%, respectively.

To this end, initial findings indicate that differences in controlled attention and working memory capacity might affect the effectiveness of CAPTCHA challenges since users of the intermediate and enhanced user groups needed less attempts than the ones of the limited user groups [35]. Such a result might be based on the fact that enhanced controlled attention and working memory capacity is needed to effectively focus a person's attention on the distorted characters among the added noise of current text-recognition CAPTCHAs.

CAPTCHA Solving Effectiveness based on Number of Refreshes. Given that the CAPTCHA mechanism also contained a refresh button that initialized the CAPTCHA with a new sequence of characters in case the characters could not be recognized by the user, the solving effectiveness of CAPTCHA could also be inferred through the number of refreshes each user made in each session. Table 6 summarizes the means of refreshes across all three levels of CAPTCHA complexity per cognitive processing group (i.e., SP, CA, WMC groups). Given that the majority of sessions (72.8%) did not include a refresh event, our analysis includes sessions that needed at least one refresh event. Shapiro-Wilk tests revealed that these distributions do not follow the normal distribution.

Table 6. Means of Refreshes per User Group

	Speed of Processing		**Controlled Attention**		**Working Memory Capacity**	
	Mean	**Std. Dev.**	**Mean**	**Std. Dev.**	**Mean**	**Std. Dev.**
Enhanced	1.22	0.42	1.38	0.65	1.14	0.37
Intermediate	1.13	0.35	1.17	0.40	1.40	0.69
Limited	1.64	0.84	1.38	0.66	1.39	0.65

On average, users with limited processing abilities initiated more refresh events than the other two user groups, however, these differences were not significant as the Kruskal-Wallis test revealed (*SOP:* $H(2)=3.571$, $p=0.168$; *CA:* $H(2)=0.503$, $p=0.778$; *WMC:* $H(2)=0.830$, $p=0.660$). In this respect, no safe conclusions can be drawn at this point in time whether cognitive processing abilities have a main effect on CAPTCHA effectiveness in terms of refresh rates. This may be due to the fact that the number of initiated refresh events was limited. Nevertheless, as the sample increased, users with limited cognitive processing abilities tended to initiate more refresh events than users with intermediate and enhanced cognitive processing abilities. In this respect, further studies need to be conducted with a greater sample over a longer period of user interactions in order to reach to more concrete conclusions about the effect of cognitive processing abilities on the effectiveness of CAPTCHA in terms of refresh events.

3.5 Validity and Limitations of the Study

The validity of a study is primarily affected by its internal, external, and ecological validity. Internal validity reflects the accuracy of data and the conclusions drawn based on this data, external validity indicates whether the data and the conclusions drawn can be generalized to a wider extend [31], and ecological validity requires that the experimental design, procedure and setting of the study must approximate the real-life context that is under investigation [32].

With the aim to increase internal validity we recruited a sample of participants already familiarized with CAPTCHAs prior to the study. Thus, the participants involved rather experienced and average than novice users with respect to CAPTCHA and therefore, the research design was setup in order to avoid inference errors. There has also been an effort to increase ecological 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. In addition, participants were required to solve the CAPTCHA challenges as a secondary task throughout the semester during real-life tasks (the primary task was to post comments on the online blogging tool). Finally, given that future studies will contribute to the external validity of the reported research, we argue that providing personalized CAPTCHA mechanisms, adapted to users' cognitive processing abilities, as well as other individual characteristics [3, 9, 33, 34, 35] could improve the overall user experience with regard to CAPTCHA tasks.

The limitations of the reported study are related to the fact that participants were only undergraduate students with an age between 17 to 26 years. Furthermore, a single assessment of users' cognitive factors through the psychometric tests might not fully justify the users' cognitive classification since other factors (e.g., emotions, urgency, etc.) might influence the users' interactions with the test. 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.

4 Conclusions

The research reported aimed to increase our understanding on supporting usable CAPTCHA designs through user modeling, and adaptivity for assisting users to accomplish efficiently and effectively usable CAPTCHA tasks. Accordingly, a three-month ecological valid user study was designed which entailed credible psychometric-based tests for eliciting the users' cognitive characteristics and a text-recognition CAPTCHA, with the aim to investigate whether individuals with different cognitive processing abilities perform differently in terms of efficiency and effectiveness in text-recognition CAPTCHA challenges with different levels of complexity.

Preliminary results reveal that cognitive processing abilities of individuals primarily affect solving efficiency of text-recognition CAPTCHA mechanisms. In particular, results demonstrated that users with enhanced controlled attention and speed of processing performed significantly faster than users with limited processing abilities across all three levels of CAPTCHA complexity. Furthermore, users with limited and intermediate cognitive processing abilities performed similarly in the low and medium complexity CAPTCHA, however a decreased performance of these users was observed in the high complex CAPTCHA type. Such a result suggests that medium complexity CAPTCHA should be provided to users with limited and intermediate cognitive processing abilities since no significant differences were observed with the lower complex CAPTCHA. Such a recommendation would increase security (medium complexity instead of low) and preserve usability in terms of efficiency at equal levels. In addition, since users with enhanced cognitive processing abilities performed slightly slower in the high complex CAPTCHA compared to the less complex CAPTCHAs of the limited and intermediate user groups, a highly complex CAPTCHA could be used in the case of users with enhanced cognitive processing abilities, thus increasing even more the security of CAPTCHA at a minimum cost of usability.

Regarding the working memory dimension, significant differences in performance have been observed between users with enhanced and limited working memory capacity (primarily in the cases of medium and high complexity CAPTCHA) suggesting that users with limited working memory capacity should be provided with less complex CAPTCHA for increasing usability. Furthermore, results revealed that speed of processing abilities might affect performance in regard with working memory capacity since a within-group analysis of enhanced WMC users revealed that the majority of them had limited and intermediate levels of processing speed, thus providing a possible explanation to the fact that enhanced WMC users performed with no significant differences in the low complex CAPTCHA, compared to limited WMC users.

Regarding effectiveness (total number of attempts), initial findings indicate that controlled attention and working memory capacity affect the effectiveness of CAPTCHA challenges since users with limited cognitive processing abilities needed more attempts than the other two user groups. On the other hand, no main effect of speed of processing on the effectiveness of solving CAPTCHAs has been revealed since the differences among user groups were not significant. This result further strengthens the validity of our theoretical background since speed of processing primarily affects efficiency of interaction, whereas in the case of effectiveness, enhanced controlled attention and working memory capacity are primarily needed to effectively focus a person's attention on the distorted characters among the added noise of CAPTCHAs. Furthermore, regarding effectiveness in terms of total number of refresh events, at this point in time, no safe conclusions can be drawn whether there is a thorough significant relationship between users' cognitive processing abilities and effectiveness in regard with CAPTCHA mechanisms since the majority of sessions did not include refresh events. Nevertheless, the analysis indicates that users with limited cognitive processing abilities initiated more refresh events to solve the CAPTCHA challenges compared to users with intermediate and enhanced cognitive processing abilities.

To this end, recent research [3, 9, 33, 34, 35] suggests that an effective CAPTCHA solution should embrace both security and usability aspects as its purpose is to provide safety of operation to Web application providers but as well usability and transparency to its end users, aiming to minimize the added cognitive effort of a casual user interacting with it. Both security and usability are important, and every CAPTCHA solution should be a balancing act between the two, aiming to achieve maximum security, but at a minimum cost to usability. Based on the presented results we suggest that following a user-centered design approach, it is necessary that designers of CAPTCHA mechanisms should clearly bear in mind individual differences of users while interacting with the system, and accordingly provide a balanced usable and secure solution. In sum, the findings of the study suggest that all three cognitive factors (speed of processing, controlled attention, working memory capacity) should be taken into consideration when designing CAPTCHAs for improving the efficiency of interactions since significant differences in all three cognitive-based user groups where observed in CAPTCHAs with different complexity. With the aim to improve effectiveness of CAPTCHA interactions, the results suggest considering controlled attention and working memory capacity of users since significant differences have been observed in CAPTCHA interactions within these cognitive-based user groups.

A practical implication of this work would be to allow users explicitly declare their preferred CAPTCHA mechanism in their Web browser preference settings, and this information would be further utilized by the Web browser to present a personalized CAPTCHA challenge. Such a scenario assumes that the Web browser and the CAPTCHA mechanism of the service provider communicate under a common standard and protocol in order to personalize the CAPTCHA challenge according to the user profile stored on the Web browser. A more sophisticated approach could be based on a recommendation engine that would implicitly present the “best-fit” CAPTCHA mechanism based on historical usage data of the user in regard with efficiency and effectiveness of CAPTCHA tasks. In this respect, the high-level research goal of such an attempt would focus around two main issues; appropriate user modeling dealing with what information is important to be incorporated in the user model and how it can be represented and extracted, and appropriate adaptation procedures dealing with what adaptation types and mechanisms are most effective to be performed and how they can be translated into adaptive user interface designs in order to improve the system’s usability and to provide a positive user experience with regard to CAPTCHA mechanisms.

Studies like the reported one can be useful for improving usable security on the World Wide Web through adaptivity in user interface designs with regard to CAPTCHA mechanisms, aiming to organize and present information and functionalities related with CAPTCHA tasks in an adaptive format to diverse user groups, by using different levels of abstractions through appropriate interaction styles, terminology, information presentation and user modeling techniques.

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).

5 References

1. von Ahn, L., Blum, M., Langford, J.: Telling Humans and Computers Apart Automatically. J. Communications of the ACM 47, 56-60 (2004)
2. von Ahn, L., Blum, M., Hopper, N.J., Langford, J.: CAPTCHA: Using hard AI problems for security. In: 22nd International Conference on Theory and Applications of Cryptographic Techniques, pp. 294-311. Springer, Heidelberg (2003)
3. Bursztein, E., Bethard, S., Fabry, C., Mitchell, J.C., Jurafsky, D.: How Good are Humans at Solving CAPTCHAs? A Large Scale Evaluation. In: IEEE Symposium on Security and Privacy, pp. 399-413. IEEE Computer Society, Washington, DC (2010)
4. von Ahn, L., Maurer, B., McMillen, C., Abraham, D., Blum, M.: reCAPTCHA: Human-Based Character Recognition via Web Security Measures. J. Science 321, 5895, 1465-1468 (2008)
5. Pogue, D.: Time to Kill Off Captchas - How the Bot-proofing of the Internet is Bringing Humans Down. J. Scientific American (2012)
6. Yan, J., El Ahmad, A.S.: A Low-cost Attack on a Microsoft CAPTCHA. In: 15th ACM International Conference on Computer and Communications Security, pp. 543-554. ACM Press, New York, NY (2008)
7. Baecher, P., Buscher, N., Fischlin, M., Milde, B.: Breaking reCAPTCHA: A Holistic Approach via Shape Recognition. In: Camenisch, J., Fischer-Hbner, S., Murayama, Y., Portmann, A., Rieder, C. (eds.) Future Challenges in Security and Privacy for Academia and Industry. LNCS, vol. 354, pp. 56-67. Springer, Heidelberg (2011)
8. Yan, J., El Ahmad, A.S.: Usability of CAPTCHAs or Usability Issues in CAPTCHA Design. In: ACM International Symposium on Usable Security and Privacy, pp. 44-52. ACM Press, New York (2008)
9. Fidas, C., Voyiatzis, A., Avouris, N.: On the Necessity of User-friendly CAPTCHA. In: ACM International Conference on Human Factors in Computing Systems, pp. 2623-2626. ACM Press, New York, NY (2011)
10. Brusilovsky, P., Kobsa, A., Nejdl, W.: The Adaptive Web: Methods and Strategies of Web Personalization. Springer, Heidelberg (2007)
11. Germanakos, P., Tsianos, N., Lekkas, Z., Mourlas, C., Belk, M., Samaras, G.: Towards an Adaptive and Personalized Web Interaction using Human Factors. In: Angelides, M., Mylonas, P., Wallace, M. (eds.) Advances in Semantic Media Adaptation and Personalization, Taylor & Francis, pp. 247-282 (2009)
12. Triantafillou, E., Pomportsis, A., Demetriadis, S., Georgiadou, E.: The Value of Adaptivity based on Cognitive Style: an Empirical Study. J. Educational Technology 35, 95-106 (2004)
13. Graf, S., Kinshuk: Advanced Adaptivity in Learning Management Systems by Considering Learning Styles. In: International Workshop on Social and Personal Computing for Web-Supported Learning Communities, pp. 235-238 (2009)
14. Riding, R.: Cognitive Style Analysis - Research Administration. Learning and Training Technology (2001)
15. Witkin, H.A., Moore, C.A., Goodenough, D.R., Cox, P.W.: Field-dependent and field-independent Cognitive Styles and their Educational Implications. J. Review of Educational Research 47, 1-64 (1977)
16. 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)
17. Hunt, E.B.: Human Intelligence. Cambridge University Press, New York (2011)

18. Demetriou, A., Spanoudis, G., Mouyi, A.: Educating the Developing Mind: Towards an Overarching Paradigm. J. Educational Psychology Review 23, 4, 601-663 (2011)
19. MacLeod, C.M.: Half a Century of Research on the Stroop Effect: An Integrative review. J. Psychological Bulletin 109, 163-203 (1991)
20. Posner, M.I., Raicle, M.E.: Images of Mind. Scientific American Library, New York (1997)
21. Stroop, J.R. Studies of Interference in Serial Verbal Reactions. J. Experimental Psychology 18, 643-662 (1935)
22. Baddeley, A.: Working Memory: Theories, Models, and Controversies. J. Annual Review of Psychology 63, 1-29 (2012)
23. Baddeley, A.: Working Memory. J. Science 255, 5044, 556-559 (1992)
24. Conway, A.R.A., Cowan, N., Bunting, M.F., Therriault, D.J., Minkoff, S.R.: A Latent Variable Analysis of Working Memory Capacity, Short-term Memory Capacity, Processing Speed, and General Fluid Intelligence. J. Intelligence 30, 163-183 (2002)
25. Shipstead, Z., Broadway, J.: Individual Differences in Working Memory Capacity and the Stroop Effect: Do High Spans Block the Words? J. Learning and Individual Differences (in press)
26. Demetriou, A., Christou, C., Spanoudis, G., Platsidou, M.: The Development of Mental Processing: Efficiency, Working Memory and Thinking. Monographs of the Society for Research in Child Development 67, 1 (2002)
27. Unsworth, N., Spillers, G.: Working Memory Capacity: Attention Control, Secondary Memory, or Both? A Direct Test of the Dual-component Model. J. Memory and Language 62, 392-406 (2010)
28. Cowan, N.: The Magical Mystery Four: How is Working Memory Capacity Limited, and Why? J. Current Directions in Psychological Science 19, 1, 51–57 (2010)
29. Klingberg, T.: The Overflowing Brain: Information Overload and the Limits of Working Memory. Oxford University Press, New York, NY (2009)
30. 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)
31. Cook, T., Campbell, D.: Quasi-experimentation: Design and Analysis Issues for Field Settings. Houghton Mifflin Company, Boston, MA (1979)
32. Brewer, M.: Research Design and Issues of Validity. In: Reis, H., Judd, C. (eds.) Handbook of Research Methods in Social and Personality Psychology. Cambridge University Press, Cambridge, pp. 3-16 (2000)
33. Belk, M., Fidas, C., Germanakos, P., Samaras, G.: Do Cognitive Styles of Users Affect Preference and Performance Related to CAPTCHA Challenges?. In: ACM International Conference on Human Factors in Computing Systems, pp. 1487-1492. ACM Press, New York (2012)
34. Holzinger, A., Lugmayr, A., Bogner, M.S.: Security and Usability Workshop. In: IEEE International Conference on Availability, Reliability and Security, IEEE Computing Society, Washington, DC (2009)
35. Belk, M., Germanakos, P., Fidas, C., Spanoudis, G., Samaras, G.: Studying the Effect of Human Cognition on Text- and Image-recognition CAPTCHA Mechanisms. In: International Conference on Human-Computer Interaction - HCI International, LNCS, Springer, Heidelberg (2013) (in press)