Int. J. Human-Computer Studies 84 (2015) 1–18

Contents lists available at ScienceDirect

Int. J. Human-Computer Studies

journal homepage: www.elsevier.com/locate/ijhcs

Do human cognitive differences in information processing affect preference and performance of CAPTCHA? ☆

Marios Belk [a,*], Christos Fidas [b,c], Panagiotis Germanakos [a,d], George Samaras [a]

[a] Department of Computer Science, University of Cyprus, Cy-1678 Nicosia, Cyprus
[b] Department of Cultural Heritage Management and New Technologies, University of Patras, Gr-26504 Patras, Greece
[c] Interactive Technologies Lab, Hci Group, Electrical and Computer Engineering Department, University of Patras, Gr-26504 Patras, Greece
[d] Suite Engineering UX, Products and Innovation, SAP SE, Dietmar-Hopp-Allee 16, 69190 Walldorf, Germany

ARTICLE INFO

Article history:
Received 2 November 2014
Received in revised form
2 June 2015
Accepted 3 July 2015
Communicated by Karen Renaud
Available online 13 July 2015

Keywords:
Human Interaction Proofs
CAPTCHA
Cognitive styles
Cognitive processing abilities
User study

ABSTRACT

A Completely Automated Public Turing test to tell Computers and Humans Apart (CAPTCHA) is a widely used security defense mechanism that is utilized by service providers to determine whether the entity interacting with their system is a human and not a malicious agent. Common design practices of current CAPTCHA schemes barely take into account cultural, contextual, and individual cognitive characteristics and abilities of users. Motivated by recent research which underpins the necessity for designing more user-friendly CAPTCHA, this paper investigates the effect of users' cognitive styles and cognitive processing abilities towards preference and task performance of CAPTCHA challenges. In the frame of the reported research, two user studies were conducted. The first study ($n=131$) explored the effect of users' cognitive styles (Verbal/Imager) on user preference and task performance of two complementary types of CAPTCHA mechanisms; text-recognition and image-recognition. The second study ($n=125$) explored the effect of users' cognitive processing abilities (speed of processing, controlled attention, working memory capacity) on task performance in regards with different levels of complexity of both text-recognition and image-recognition CAPTCHA. Analysis of results revealed interaction effects of users' cognitive processing characteristics towards preference and performance of CAPTCHA, suggesting that individual differences at such an intrinsic level are important to be considered for designing more usable and user-centric CAPTCHA challenges.

1. Introduction

Human Interaction Proofs (HIP) are wide spread security defense mechanisms for constructing a high-confidence proof that the entity interacting with a remote service is a human being, and not malicious software (Chellapilla et al., 2005). A Completely Automated Public Turing test to tell Computers and Humans Apart (CAPTCHA) (von Ahn et al., 2004) is a HIP challenge-response test widely used today to protect Web applications, services and interfaces against automated software agents whose purpose is to degrade the quality of a provided service. CAPTCHA mechanisms commonly require users to respond to visual cognitive-based challenges (e.g., recognize and type characters that are illustrated in a distorted form on the screen) or audio-based challenges for individuals with vision problems. Such challenges are based on the assumption that they can be easily solved by humans but present significant difficulty for computing systems.

Designing a CAPTCHA mechanism is an inevitable balancing act between usability and security. Increasing the complexity of a CAPTCHA challenge (e.g., by increasing the distortion of characters or increasing the number of images), increases the security of the mechanism, but significantly decreases its usability (Bursztein et al., 2011, 2014; Golle, 2008). Recently, a high number of research works underpinned the necessity for designing user-friendly CAPTCHA since several studies revealed that requiring users to solve CAPTCHA challenges is a difficult and demanding task that decreases the overall user experience with an interactive system (Fidas et al., 2011; Bursztein et al., 2010; Yan and El Ahmad, 2008). Such tasks are known to add a considerable cognitive burden to users (Fidas et al., 2011). In addition, since these challenges interrupt the users' primary task of interaction, users might not be able to complete or even abandon their task with a system. From an accessibility perspective, studies have shown that visual CAPTCHA offer little support for users with vision problems (Bigham and Cavender, 2009; Holman et al., 2007).

☆This paper has been recommended for acceptance by Karen Renaud.

* Correspondence to: Department of Computer Science, University of Cyprus, PO Box 20537, 1678 Nicosia, Cyprus. Tel.: +357 22 892700; fax: +357 22 892701.
E-mail addresses: belk@cs.ucy.ac.cy (M. Belk), fidas@upatras.gr (C. Fidas), panagiotis.germanakos@sap.com (P. Germanakos), cssamara@cs.ucy.ac.cy (G. Samaras).

http://dx.doi.org/10.1016/j.ijhcs.2015.07.002
1071-5819/

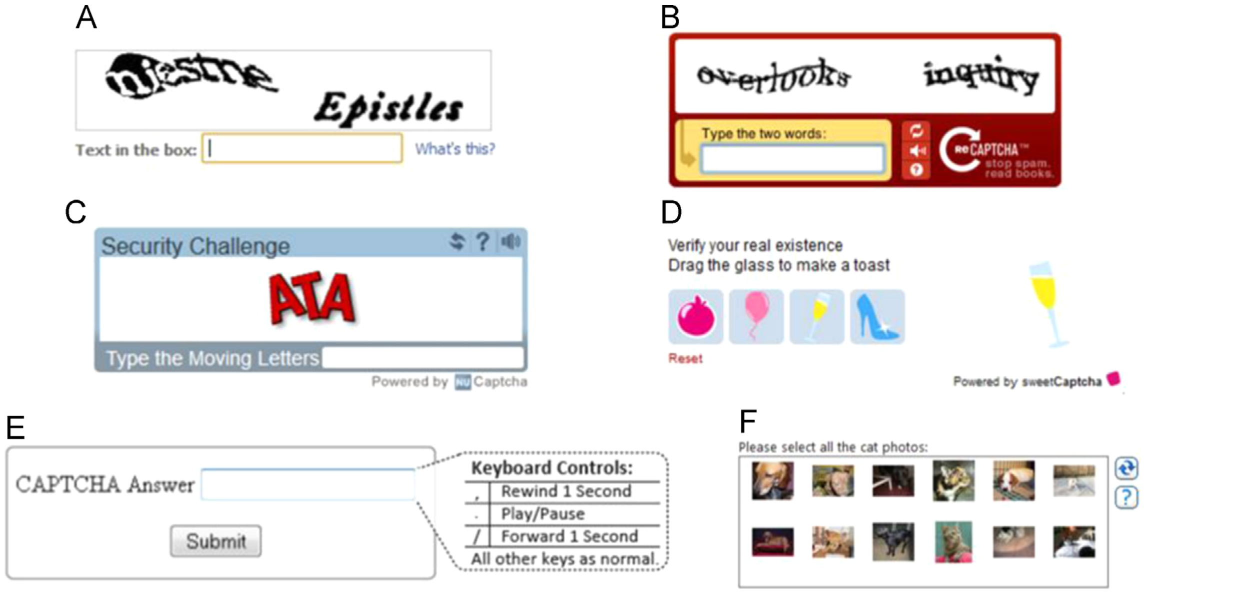


Fig. 1. (A) Facebook CAPTCHA (text-recognition), (B) reCAPTCHA (text- and speech-recognition), (C) NuCAPTCHA (animated text-recognition), (D) SweetCAPTCHA (drag-and-drop interaction), (E) non-visual Access CAPTCHA (speech-recognition), (F) Microsoft ASIRRA (image-recognition).

In this realm, aiming to improve the user experience during such interactions, but at the same time preserve security of applications and services, researchers promote different visual and interaction designs of CAPTCHA challenges (see Moradi and Keyvanpour, 2014 for a recent review). Current CAPTCHA implementations can be classified into three broad categories: *text-recognition*, *image-recognition*, and *speech-recognition*. Fig. 1 illustrates some noteworthy CAPTCHA implementations of each category.

Text-recognition CAPTCHA mechanisms are currently the most widely used (Bursztein et al., 2010, 2014) and require from a legitimate user to type alphanumeric characters based on a distorted image that appears on the screen. Popular text-recognition CAPTCHA include among others reCAPTCHA (von Ahn et al., 2008), Google CAPTCHA (Bursztein et al., 2014) and BaffleText (Chew and Baird, 2003). *Image-recognition CAPTCHA mechanisms* are usually based on image puzzle problems and annotation of static and animated images. For example, in ASIRRA (Elson et al., 2007) users are required to select pictures that illustrate cats among dogs. SEMAGE (Vikram et al., 2011) similarly requires users to recognize the content of a set of images, but as well understand and identify the semantic relationship between a subset of them. Another popular example includes What's Up CAPTCHA (Gossweiler et al., 2009) that requires from users to adjust randomly rotated images to their upright orientation. *Speech-recognition CAPTCHA mechanisms* are usually based on audio comprehension which principally require users to enter alphanumeric characters listened from a recording of a combination of simple words and numbers where disturbance and noise has also been added. Speech-recognition CAPTCHA are more difficult to solve and internationalize, and more demanding in terms of time and efforts compared to text-recognition and image-recognition CAPTCHA (Bigham and Cavender, 2009; Bursztein et al., 2010). Nevertheless, speech-recognition CAPTCHA have become an alternative for visually-impaired people that aim to improve usability and allow easy access to users (Davidson et al., 2014; Bigham and Cavender, 2009; Holman et al., 2007; Gao et al., 2010).

The literature also reveals a high number of alternative CAPTCHA mechanisms that follow hybrid approaches that combine text- and image-recognition challenges, drag-and-drop interactions, semantic approaches, etc. Examples include NuCAPTCHA that illustrates animated instead of static text in the challenge (NuCAPTCHA Inc., 2015), Emerging CAPTCHA (Xu et al., 2014) which is an alternative approach that addresses security flaws found in NuCAPTCHA, Solve-Media CAPTCHA that incorporates brand advertisements in the challenge and users are required to type the text of an advertiser's brand text (SolveMedia, 2015), video approaches such as the work of Kluever and Zanibbi (2009) that proposed a technique for using content-based video labeling as a CAPTCHA challenge and users are then required to label these videos to pass the challenge, and SweetCAPTCHA (2015) that is an action-based CAPTCHA in which users are required to drag-and-drop specific objects onto other objects (e.g., *"drag the shoes into the box"* or *"drag the glass to make a toast"* (Fig. 1D)).

A common practice with regards to the aforementioned visual and interaction designs of CAPTCHA mechanisms is that they do not primarily take into consideration the individual characteristics of users but rather follow a one-size-fits-all paradigm, i.e., the visual and interaction design of CAPTCHA is rarely personalized to the individual characteristics of users (e.g., cognitive processing abilities). Nevertheless, recent research revealed that individual differences have a main effect on task performance and user preference of CAPTCHA (Fidas and Voyiatzis, 2013; Belk et al., 2012; Wei et al., 2012; Albert et al., 2010; Banday and Shah, 2011), suggesting that user-adaptive and personalized CAPTCHA mechanisms could improve the user experience and user acceptance of CAPTCHA. Consequently, an important step toward designing personalized and user-centric CAPTCHA mechanisms is to identify which individual characteristics are considered important enough and might affect users' interactions with such security mechanisms.

In this context, bearing in mind that solving a CAPTCHA challenge (text, image, sound) is primarily a human cognitive processing task; users are required to process and recognize textual, graphical or audio information, we suggest that human cognitive differences in information processing should be investigated and integrated in the user interface design process of CAPTCHA challenges. Accordingly, this paper reports two subsequent user studies that aim to further understand human-computer interactions in such realms and investigate the effect of users' cognitive processing characteristics on preference and task performance of different designs of CAPTCHA challenges. The first study ($n=131$) explores the effect of users' cognitive styles (Verbal/Imager) on user preference and task performance of two complementary types of CAPTCHA mechanisms; text-recognition and image-recognition. The second study ($n=125$) explores the effect of users' cognitive processing abilities (speed of processing, controlled attention, working memory capacity) on task performance in regards with different levels of complexity (in terms of added noise and distortion on the characters/images) of both text-recognition and image-recognition CAPTCHA challenges.

The rest of the paper is organized as follows: next we analyze the underlying theory of this work. Then, we describe the method and the design of two user studies, and subsequently we analyze and discuss the findings of the studies. Finally, we present the practical implications, the validity and the limitations of this work and conclude the paper with future research prospects.

2. Individual differences in cognitive processing styles and abilities

The theoretical background of this work is primarily based on theories of individual differences in cognitive processing styles and abilities (Demetriou et al., 2013; Riding and Cheema, 1991; Peterson et al., 2009; Kozhevnikov, 2007), suggesting that individuals have preferred ways of representing and processing textual and graphical information, as well have different cognitive abilities in processing information (e.g., in terms of speed).

A number of researchers have focused on high-level cognitive processes such as *cognitive styles*, which explain empirically observed differences in mental representation and processing of information (Riding and Cheema, 1991; Peterson et al., 2009; Kozhevnikov, 2007). A particularly important cognitive style is the Verbal/Imager dimension that refers to how individuals process information and indicates their preference for representing information in words (Verbals), or in mental pictures (Imagers) (Riding and Cheema, 1991). *Verbals* represent the information they read, see or listen in words or verbal associations. Individuals being Verbals prefer and perform more efficiently when hypermedia content is presented in the form of text. Verbals also have great reading accuracy and are better at recalling acoustically complex and unfamiliar text (Liu and Ginther, 1999). *Imagers* represent information in mental pictures, focus their attention internally and tend to be passive which means they are primarily triggered by their thoughts, memories, etc. (Liu and Ginther, 1999). Imagers prefer and perform more efficiently when the hypermedia content is provided in the combination of graphical and textual representation, but do not perform efficiently when an exclusively verbal representation is provided (Ghinea and Chen, 2008).

Researchers have also attempted to explain the functioning of the human mind in terms of more elementary cognitive processes. These include the *speed of processing*, which refers to the maximum speed a given mental act may be efficiently executed (MacLeod, 1991); *controlled attention*, which refers to cognitive processes to identify and concentrate on goal-relevant information and inhibit attention to irrelevant stimuli (Stroop, 1935; MacLeod, 1991); and *working memory capacity*, which is defined as the maximum amount of information the mind can efficiently activate during information processing (Baddeley, 2012, 1992). Studies revealed the relationship between speed of processing, controlled attention and working memory capacity (Conway et al., 2002; Shipstead and Broadway, 2013; Polderman et al., 2006). For example, enhanced speed of processing facilitates access to information that is sustained in the working memory system (Baddeley, 1992) as well as enables individuals to handle more efficiently information flow during problem solving (Hale and Fry, 2000). Also, individuals with enhanced working memory capacity are less susceptible to the attention interference since they have improved controlled attention abilities (Unsworth and Spillers, 2010; Conway et al., 2001).

Various research works argue that the aforementioned cognitive processing characteristics have an effect on comprehension, learning and problem solving (Demetriou et al., 2013; Shipstead and Broadway, 2013; Unsworth and Spillers, 2010). 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 recognizing the distorted text of a CAPTCHA challenge. Accordingly, various research attempts have been reported that correlate cognitive processing characteristics of users on issues related to task efficiency and effectiveness. A number of studies confirmed that cognitive factors affect search and browsing behavior of users in interactive systems (Chen and Liu, 2008; Belk et al., 2013), learning performance and comprehension of educational and hypermedia environments (Germanakos et al., 2008; Ghinea and Chen, 2008; Papanikolaou et al., 2006), eye-gaze behavior within information visualization (Toker et al., 2013), perceived distributed multimedia quality (Chen et al., 2006), and task performance of troubleshooting diagnosis systems (Cegarra and Hoc, 2006). On the contrary, various studies concluded that cognitive factors do not have a main effect on users' task performance and preference within hypermedia environments (Mitchell et al., 2004; Brown et al., 2006).

3. Method of study

In this section we describe the method and design of two user studies that aimed to investigate the effects of users' cognitive differences in information processing, on preference and task performance related to different designs of CAPTCHA.

3.1. Research questions

Research on CAPTCHA mechanisms has become a complex endeavor since it embraces several parameters (human and design specific) that need to be taken into account. Thus, there is a need for solid research frameworks which will assist in understanding human interactions in such settings. In this context, the question remains whether and how state of the art socio-cognitive theories can be adopted as an analysis framework aiming to assist the design of more user-centric and usable CAPTCHA mechanisms. Motivated by the aforementioned rational, this work aims to contribute toward this direction by investigating: (i) whether there is an observable main effect of users having a particular style of representing and processing information cognitively (verbal or visual), on user preference and task performance of two different CAPTCHA designs (textual or graphical); and (ii) whether there is an observable main effect of users having different cognitive processing abilities (e.g., enhanced speed of processing and high levels of working memory capacity), on different levels of CAPTCHA challenge complexity (e.g., number of characters/images, distortion of characters/images, etc.).

Accordingly, the following main research questions are investigated:

- Are there significant differences between two CAPTCHA designs (text-recognition and image-recognition) regarding user preference, task efficiency and success rate among users with different cognitive styles?
- Are there significant differences between various levels of visual complexity in text-recognition and image-recognition

Fig. 2. Text-recognition CAPTCHA (left) and image-recognition CAPTCHA (right).

CAPTCHA regarding task efficiency and success rate among users with different cognitive processing abilities?

3.2. Sampling and procedure

Two subsequent user studies were conducted for the purpose of this research. In the first user study, a total of 131 undergraduate Computer Science students participated (76 male, 55 female, age 20–25, mean 23) having recorded the same number of CAPTCHA sessions. Similarly, in the frame of the second user study, a total of 125 undergraduate Electrical Engineering students participated (48 male, 77 female, age 18–28, mean 20). A total of 125 CAPTCHA sessions have been recorded.

A series of accredited Web-based psychometric instruments were developed that highlight differences in cognitive styles and cognitive processing abilities. In addition, two types of CAPTCHA mechanisms were developed; a *text-recognition CAPTCHA* that requires from users to recognize and enter distorted alphanumeric characters, and an *image-recognition CAPTCHA* that requires from users to recognize and select specific images of a particular theme (Fig. 2).

The developed psychometric instruments and CAPTCHA mechanisms were applied in online University courses in the context of two user studies. Both user studies followed a two-phase methodological approach that entailed a cognitive factor elicitation phase for highlighting the participants' cognitive characteristics, and a user interaction phase with the developed CAPTCHA mechanisms in which users solved a CAPTCHA challenge.

Next we present in detail the developed CAPTCHA mechanisms and the psychometric instruments that were used in two experimental studies.

3.3. CAPTCHA mechanisms used in the studies

By following state of the art literature in CAPTCHA, at a first level we chose to investigate traditional text-recognition and image-recognition CAPTCHA mechanisms since these are currently the most widely researched and applied CAPTCHA scheme categories (Bursztein et al., 2010; Zhu et al., 2010; Moradi and Keyvanpour, 2014). In particular, the choice was based on the fact that solving each CAPTCHA challenge (text- and image-recognition) requires processing and recognition of text or images in which users utilize their verbal and image cognitive sub-systems that are related to the Verbal/Imager cognitive style theory referred in this work (Riding and Cheema, 1991). We intentionally did not investigate hybrid or alternative approaches at this stage (e.g., drag-and-drop interactions, semantic approaches, etc.) since we aimed to isolate and control the type of content illustrated in the challenge (textual or graphical) so that users would utilize their verbal and image cognitive sub-systems while processing information and solving the challenge. Furthermore, a speech-recognition mechanism was also intentionally not investigated in this work since it is considered a significantly more demanding task in terms of solving time and is mostly used for users with physical impairments (Bigham and Cavender, 2009).

In the analysis that follows, we describe both CAPTCHA mechanisms that were utilized focusing on their respective design and development choices, security metrics and visual design.

3.3.1. Text-recognition CAPTCHA mechanism

A traditional text-recognition CAPTCHA mechanism was developed that requires from users to recognize and enter the correct sequence of alphanumeric characters that are illustrated in a distorted form on the screen. The CAPTCHA mechanism also includes a refresh button that initializes a challenge by reloading a new set of characters. The development of the text-recognition CAPTCHA mechanism was based on a similar technical and architectural approach followed by reCAPTCHA (von Ahn et al., 2008) as well as freely available open-source software (Securimage v.3.5.2, 2014). We intentionally chose an existing open-source CAPTCHA mechanism that would be extensible and customizable in order to design and customize the text features (font size, length, color, rotation etc.) for the purpose and aim of each user study. This enabled us to control the design factors and security metrics of the challenges utilized in the studies (e.g., keep the visual complexity at the same level for all participants, or adjust the complexity level for specific groups of participants). The design and customization of the text features illustrated in the CAPTCHA challenge was based on design and security guidelines proposed in Bursztein et al. (2011).

3.3.2. Image-recognition CAPTCHA mechanism

The development of the image-recognition CAPTCHA mechanism was based on Microsoft ASIRRA which presents to the users 12 images (40 pixel $\times$ 40 pixel per image) illustrating cats and dogs requiring from them to recognize and select the images that display cats (Elson et al., 2007). The same interaction design was used as the one provided by ASIRRA in which users need to hover over each image in order to view the larger version of the image. The CAPTCHA mechanism also includes a refresh button that initializes a challenge by reloading a new set of images. Among a high number of existing image-recognition CAPTCHA mechanisms (Zhu et al., 2010), the choice of ASIRRA was based on the following reasons: (i) solving an image-based challenge in ASIRRA primarily entails a visual search task (that is affected by cognitive styles (Angeli et al., 2009)) in which users are required to search, find and recognize images of a particular theme by utilizing their image cognitive sub-system. Thus we aimed to isolate this human cognitive processing task and investigate whether Verbal and Imager cognitive styles affect the preference and task performance of this particular challenge; (ii) ASIRRA belongs to a broad image-recognition CAPTCHA category (the distinguishing CAPTCHAs) which is among the three main image-recognition CAPTCHA scheme categories (i.e., naming images, distinguishing images, identifying anomalies) (Chew and Tygar, 2004); (iii) ASIRRA is a highly cited and considered a representative image-recognition

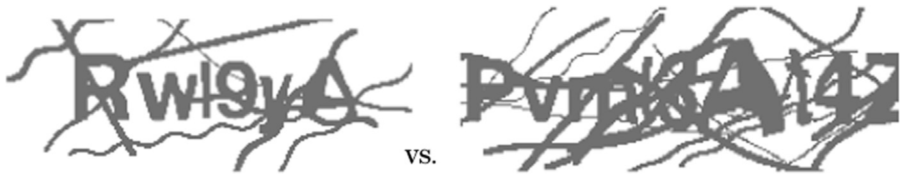


Fig. 3. Baseline vs. higher complexity text-recognition CAPTCHA.

CAPTCHA scheme (Zhu et al., 2010); and (iv) according to various usability evaluations and CAPTCHA reviews (Elson et al., 2007; Zhu et al., 2010), the mean task completion time is more efficient and solving accuracy is higher in contrast to other image-recognition CAPTCHA mechanisms that exist in the literature (Zhu et al., 2010).

3.3.3. *CAPTCHA security metrics*

The success rate of an attack is the primary metric to evaluate CAPTCHA attack effectiveness (Zhu et al., 2010). Strictly, attackers should not have a success rate higher than 0.01%; automated scripts should not be able to successfully solve more than 1 CAPTCHA challenge in 10,000 attempts (Chellapilla et al., 2005). Nevertheless, researchers have reported that such a security goal is very ambitious and challenging when designing CAPTCHA mechanisms (Bursztein et al., 2011; Zhu et al., 2010). Accordingly, the success rate of attacks are also acceptable at a value of 1% (Bursztein et al., 2011) when IP monitoring is used in combination with the CAPTCHA challenge, such as the token bucket scheme proposed in Elson et al. (2007). In essence, the token bucket scheme "punishes" users that fail to solve the challenge at first attempt by requiring them to solve two or more consecutive CAPTCHA challenges.

In this context, the security issues for the design and development of the text-recognition CAPTCHA were addressed based on design guidelines and suggestions proposed by Bursztein et al. (2011). Based on the guidelines, the success rate of an attack is estimated to be less than 1%. For the image-recognition CAPTCHA utilized in this work (ASIRRA), results reported in Elson et al. (2007) have shown that the probability of attack success is estimated to be 0.2%. On the contrary, machine learning attacks on the original version of ASIRRA, developed by Golle (2008), showed a high attack success rate (10.3%). Nevertheless, Golle (2008) suggested that with appropriate safeguards (e.g., token bucket scheme), ASIRRA "continues to offer an appealing balance between security and usability" (Golle, 2008). Specifically, attacks with the token bucket scheme enabled have revealed that the success rate of an attack on a 12-image ASIRRA challenge is estimated to be approximately 1%. This success rate value could be further decreased by including a larger number of images in the challenge as well as using greyscale images, instead of color images (Golle, 2008).

3.3.4. *CAPTCHA visual designs*

Based on the aforementioned security considerations, the security of CAPTCHA mechanisms is highly affected by the added complexity of the visual design of the CAPTCHA challenge (Bursztein et al., 2011, 2014; Zhu et al., 2010). Thus, in this work we primarily focus on the visual complexity of CAPTCHA mechanisms (that principally require users' cognitive processing of information) with the aim to investigate how users (with limited or enhanced cognitive processing abilities) are affected in terms of task completion performance. In this context, we have designed the two CAPTCHA challenges to entail different complexity levels.

The complexity levels of both text-recognition and image-recognition CAPTCHA were respectively based on the design and security guidelines suggested by Bursztein et al. (2011) and Golle (2008). According to Bursztein et al. (2011), increasing the security of text-recognition CAPTCHA could be achieved with the following design principles: (i) randomize the CAPTCHA length and font size; (ii) rotate the characters in a wave fashion; (iii) use lines with the same width and color as the characters; and (iv) collapse the characters. Using a background image with noise in the challenge has shown to be insecure and therefore we excluded this technique from our text-recognition CAPTCHA design. Regarding the image-recognition CAPTCHA, based on Golle (2008), increasing the security of the ASIRRA CAPTCHA could be achieved as follows: (i) increase the number of images in the challenge; and (ii) use greyscale instead of colored images. We intentionally did not degrade the quality of the images, nor used distortion since this is unlikely to increase the security of the particular image-recognition mechanism, but rather only decrease its usability (Golle, 2008).

Two different levels in terms of visual complexity have been designed; a design with baseline security and a higher complex design. In the case of text-recognition CAPTCHA, 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 baseline complexity CAPTCHA entailed a random number of 5–7 characters and 40% character rotation, collapsing and lines, while the higher complex CAPTCHA entailed 8–10 characters, and 60% character rotation, collapsing and lines, as illustrated in Fig. 3.

In the case of image-recognition CAPTCHA, the criteria for developing the different levels of complexity were based on the number of images illustrated in each challenge and the type of image color used (greyscale or color). The low complex CAPTCHA illustrated a 12-image challenge with colored images (same as the baseline ASIRRA CAPTCHA) while the higher complex CAPTCHA illustrated a 14-image challenge with greyscale images, as illustrated in Fig. 4.

3.4. *Cognitive factor elicitation tools used in the studies*

A number of online psychometric tests were developed. The users' Verbal/Imager cognitive styles were elicited by exploiting Riding's Cognitive Style Analysis test (CSA) (Riding, 1991; Riding and Cheema, 1991) considering also guidelines reported in Rezaei and Katz (2004). The users' cognitive processing abilities were elicited by exploiting two Stroop-like tests for eliciting the users' speed of processing and controlled attention, and two working memory capacity tests as utilized in Demetriou et al. (2013). In principal, all tests measure response times of users on specially designed aptitude tasks that require cognitive processing. Depending on the response time and the provided answer to each task, the users' cognitive characteristics are highlighted on a specific scale (e.g., Verbal-Intermediate-Imager, Limited-Enhanced cognitive processing ability). We next describe each psychometric test and how cognitive characteristics are elicited.

3.4.1. *Cognitive styles' test*

The CSA test indicates an individual's tendency to process information verbally or in mental pictures. An individual's style on the Verbal-Imager dimension is obtained by presenting a series of 48 questions about conceptual category and appearance (i.e., color) to be judged by the users to be true or false. A total of 24

Fig. 4. Baseline (colored) vs. higher (greyscale) complexity image-recognition CAPTCHA. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Study	Cognitive Factors	CAPTCHA Design Factors	Usability Factors
A	**Cognitive Styles** ▪ Verbal ▪ Imager ▪ Intermediate	**CAPTCHA Type** ▪ Text-recognition ▪ Image-recognition	▪ User Preference ▪ Completion Time ▪ Success Rate
B	**Cognitive Abilities** ▪ Limited ▪ Enhanced	**CAPTCHA Visual Complexity** ▪ Baseline ▪ Higher	▪ Completion Time ▪ Success Rate

Fig. 5. Factors that were investigated in each user study.

statements require comparing two objects conceptually (e.g., "Are ski and cricket the same type?"). The remaining 24 statements require comparing the color of two objects (e.g., "Are cream and paper the same color?"). It is assumed that Verbals respond faster than Imagers in the conceptual types of stimuli because the semantic conceptual category membership is verbally abstract in nature and cannot be represented in visual form (Riding, 1991). On the other hand, it is assumed that Imagers respond faster than Verbals in the appearance statements (color) since the objects can be represented as mental pictures and the information for the comparison can be obtained directly and rapidly from these images (Riding, 1991).

3.4.2. Speed of processing and controlled attention tests

Two Stroop-like tasks were devised to measure reaction time to address speed of processing and controlled attention. For measuring speed of processing, users are required to read a number of words denoting a color written in the same or different ink color (e.g., the word *"red"* illustrated in red ink color). For measuring controlled attention, a similar Stroop-like task is devised, but instead of denoting the written word itself, users are required to recognize the ink color of words denoting a color different than the ink (e.g., the word "green" illustrated in blue ink). In each test, a total of 18 words are illustrated to the users illustrating the words "red", "green" or "blue" either written in red, green or blue ink color. The participants are required to press the R key of the keyboard denoting the selection of "red", the G key for "green" and the B key for "blue".

3.4.3. Users' working memory capacity tests

Two tests addressed storage capacity in short-term memory; a *visual* and a *verbal test*. The *visual test* illustrates a geometric figure on the screen and the user is required to memorize the figure. Thereafter, the figure disappears and 5 similar figures are illustrated on the screen, numbered from 1 to 5. The user is required to provide the number (utilizing the keyboard) of the corresponding figure that was the same as the initial figure. The test consists of 21 trials. As the user correctly identifies the figures of each trial, the test provides more complex figures as the levels increase indicating an enhanced visual working memory capacity. The *verbal test* shows a series of statements and requires users to respond whether they are true or false. In addition, users are required to remember the last word of each sentence and then write that word. The test includes six levels of difficulty, e.g., in level 3, participants are required to respond true/false on three successive sentences and have 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 user must respectively respond *true*, *true* and *false*, and then provide the word *"sharp"*, *"shining"* and *"fur"* to the system. The total number of correct responses of each participant indicates his/her verbal working memory capacity.

3.4.4. Users' cognitive factor calculation

For each cognitive stimulus of the aforementioned tests, the response time and the provided answer are recorded. Based on the users' responses, the following two user cognitive characteristics are elicited: (i) a user is either a Verbal, either an Imager or an Intermediate, and (ii) a user has either limited or enhanced cognitive processing abilities.

Responses of the cognitive styles' test are processed as follows: the average response time of all valid and correct responses is calculated on each of the two stimuli types (verbal and imagery) of the psychometric test, and then the ratio between the average response times on the verbal and imagery stimuli is calculated. Users with a low ratio are considered Verbals, users with a high ratio are considered Imagers, while users in between the two end points are considered Intermediates (Riding, 1991; Rezaei and Katz, 2004; Belk et al., 2013).

Regarding the speed of processing and controlled attention tests, the average response time of all valid and correct responses is calculated, indicating how fast users process information, whereas for the working memory capacity tests the total number of correct responses is measured indicating the users' level of visual and verbal working memory capacity. For all the cognitive processing abilities' tests, a normalization by z-score is performed on the data, since speed of processing and controlled attention measure speed (average in seconds), whereas working memory measures capacity (total number of correct responses). The final z-value indicates a user's cognitive processing ability, with a low value indicating an enhanced cognitive processing ability of that user, and a high value indicating a limited cognitive processing ability of that user (Belk et al., 2014).

For classifying the users into specific cognitive factor groups (Verbal/Imager cognitive style and limited/enhanced cognitive processing ability), a k-means clustering algorithm (presented in Belk et al. (2014)) is applied on the users' processed responses (cognitive styles' ratio and z-value), utilizing an existing representative sample of 800 undergraduate students of the same university whose cognitive styles and cognitive processing abilities were elicited in past user studies of the authors.

3.5. Experimental design

This research embraced two independent user studies that were conducted at two consecutive time points. Fig. 5 illustrates the factors being investigated in each user study.

User Study A aimed to investigate whether individual differences in cognitive styles affect user preference and task performance, in terms of task completion time and task completion success rate, of two different types of CAPTCHA; text-recognition and image-recognition. User Study B aimed to investigate whether individual differences in cognitive processing abilities affect task completion time and task completion success rate of two different levels of CAPTCHA challenge complexity (baseline and higher) of both text-recognition and image-recognition CAPTCHA. In User Study B, the users' cognitive styles were used as control factors; based on the cognitive styles of each user (Verbal/Imager/Intermediate), the system provided a text-recognition CAPTCHA to Verbal and Intermediate users, and an image-recognition CAPTCHA to Imager users since an observable main effect of cognitive styles of users on CAPTCHA task completion efficiency had been observed in User Study A (Belk et al., 2012).

The participation in both user studies was voluntary and all individuals agreed to an online consent form before participating. In particular, participants were informed that the data they provided and their interactions with the system would be processed and used anonymously as part of an experimental user study of the researchers' group. No further details about the aim of the study, nor the type of interaction data recorded (e.g., time to complete the challenge) were provided to the participants in order to avoid bias effects. We next describe the experimental design of each user study.

3.5.1. Experimental design (A)

The first user study embraced a between-subject design, aiming to examine whether cognitive styles of users affect preference and performance (task efficiency and effectiveness) on two different types of CAPTCHA challenges; text-recognition and image-recognition. An invitation was announced on the Web-sites of various undergraduate Computer Science courses in order to recruit the participants. The aim of this selection process was to recruit a representative sample of participants that were already familiarized with CAPTCHA challenges based on the fact that Computer Science students are faced daily with CAPTCHA challenges in online courses, forums, blogs, social networking Web-sites, etc.

The participants were asked to visit a Web-page in order to take part in the study. The Web-page provided information and guidelines regarding the study as well as the two CAPTCHA challenges (text and image). The users first provided basic demographic information (age, gender). After, the users were required to choose between the two variations of CAPTCHA (i.e., text- vs. image-recognition) and then solve the preferred CAPTCHA challenge. For the purpose of the study, the complexity level of each CAPTCHA challenge was the same (i.e., 8 characters with the same percentage of noise and distortion was used in the text-recognition challenge, whereas 12 colored images were used in the image-recognition challenge). After solving the CAPTCHA challenge, the users were redirected to an online psychometric test aiming to elicit the users' cognitive styles.

3.5.2. Experimental design (B)

The second user study was applied in the frame of a registration process of a university Computer Science course as part of students' enrolment in the course's Web-site. Main aim of this process was to increase the ecological validity of the users' interactions with CAPTCHA challenges since the Web-site would be used by the students to view and download material of their course throughout the semester. The user study embraced a between-subject design, aiming to investigate whether cognitive processing abilities of users affect task completion efficiency and effectiveness of different levels of CAPTCHA complexity (i.e., baseline or higher level of complexity for both text- and image-recognition CAPTCHA).

The user study was conducted in a controlled lab setting and was split in two phases at two different time stamps: Phase A for eliciting the users' cognitive processing characteristics (cognitive styles and cognitive processing abilities), and Phase B for enrolling the users in the online course and record their interactions with the CAPTCHA challenges.

In Phase A, users initially interacted with the developed psychometric tests by providing their unique student identity, with the aim to elicit their cognitive processing characteristics. The results of each psychometric test were bound with each user's unique student identity in order to map a particular CAPTCHA type (text or image) to the user and relate his/her interactions with the CAPTCHA challenge in Phase B. For the purpose of the study, in order to proceed with Phase B, all participants interacted first with all the psychometric tests in order to perform the cluster analysis for classifying the users into cognitive factor groups based on the cognitive styles' ratios and cognitive processing abilities' z-values.

In Phase B, users enrolled in the course through a registration form by creating a username (using their student identity) and a password, and providing their age, gender and email. The registration form also included the developed CAPTCHA mechanism which required users to solve either a text-recognition or an image-recognition CAPTCHA challenge that was decided based on their cognitive styles (these were retrieved from the database according to the provided student identity). With the aim to investigate the effects of cognitive processing abilities of users on the complexity level of CAPTCHA challenges, the system randomly provided different levels of CAPTCHA complexity so that half of the users would interact with a baseline complexity CAPTCHA and the other half with a higher complex CAPTCHA. The allocation was based on the users' cognitive processing abilities so that the complexity levels were balanced across all user groups (limited and enhanced cognitive processing abilities).

3.6. Data analysis

In both studies, client-side and server-side scripts were developed for measuring the users' interactions with the system. The data captured during the user studies are grouped as follows:

User Data consists of data about the users' individual cognitive processing characteristics. In particular, based on the users' interactions with the psychometric tests, users were classified as Verbals, Imagers or Intermediates, having limited or enhanced cognitive processing abilities.

Task Data consists of data about the task performed by the users. This includes the type of CAPTCHA (text or image), the level of complexity (baseline or higher), the total time (in seconds) and total number of attempts to successfully solve the CAPTCHA challenge. Based on the total number of attempts, the success rate (in percentage) of each session was also calculated. For example, a user that solved the challenge at first attempt had a success rate of 100%, whereas a user that solved the challenge at third attempt (the first and second attempt failed and the third succeeded) had a success rate of 33%. Additional data were recorded such as the

Table 1
Variables used in the analysis.

Data categories	Variables
User data	Cognitive styles (verbal/imager/intermediate)
	Cognitive processing abilities (limited/enhanced)
Task data	Type of CAPTCHA (text/image)
	Level of complexity (baseline/higher)
	Total time to complete (s)
	Number of attempts (ordinal)
	Success rate (percentage)
	Number of refreshes (ordinal)

Table 2
Number of users per cognitive styles' group (User Study A).

Verbals	Imagers	Intermediates	Total
65	43	23	131

number of times the challenge was refreshed (in which a new challenge was reloaded).

In the analysis, the total time, number of attempts, success rate and number of refreshes are the dependent variables, whereas the users' cognitive styles, cognitive processing abilities, type of CAPTCHA and the level of complexity are the independent variables. Table 1 illustrates the variables used in the analysis.

4. Analysis of results

In this section we analyze the results based on the users' interactions with the CAPTCHA challenges. The familiarity factor of our sample in solving text-recognition challenges should be carefully considered when interpreting the results, since all participants of both studies have experience in solving rather text-recognition than image-recognition CAPTCHA challenges.

For our analysis, we separated users in three categories based on cognitive styles (Verbal/Imager/Intermediate) and in two categories based on cognitive processing abilities (limited/enhanced). Tables 2 and 3 summarize the number of users in each group, respectively, for User Study A and User Study B.

4.1. User preference related to CAPTCHA challenges (User Study A)

In the first user study, participants were asked to choose between two variations of CAPTCHA (i.e., text- vs. image-based). In Table 4 we summarize the CAPTCHA preferences according to the users' cognitive styles.

A binomial statistical test was conducted to examine whether there is a general preference relating text- or image-recognition CAPTCHA challenges (H_0: p(text-recognition)$=0.5$ and p(image-recognition)$=0.5$). The results revealed that there is significant preference towards text-recognition CAPTCHA challenges ($p<0.01$). Furthermore, a Pearson's chi-square test was conducted to examine whether there is a relationship between users' cognitive styles and their preference towards a specific type of CAPTCHA challenge (i.e., text- or image-recognition). The results revealed that there is no significant relationship between these two variables (*Chi square value*$=0.791$, $df=2$, $p=0.673$). As a consequence, no safe conclusion can be drawn at this stage whether cognitive styles of users influence their preference towards a specific type of CAPTCHA challenge.

However, examining each cognitive styles' group individually with respect to preference towards a particular CAPTCHA type, it has been identified that users of the Verbal class have significant positive preference towards text-recognition CAPTCHA (*Chi square value*$=5.554$, $df=1$, $p<0.02$). In contrast, users belonging to the Imager class (*Chi square value*$=0.581$, $df=1$, $p=0.446$) and Intermediate class (*Chi square value*$=1.087$, $df=1$, $p=0.297$) have not shown a clear preference towards one or the other direction (i.e., text- vs. image-recognition CAPTCHA challenge).

Table 3
Number of users per cognitive styles' and cognitive processing abilities' group (User Study B).

Cognitive processing abilities	Cognitive styles			
	Verbals	Imagers	Intermediates	Total
Limited	21	15	11	47
Enhanced	38	28	12	78
Total	59	43	23	125

Table 4
Users' cognitive styles vs. CAPTCHA preference.

Cognitive styles	CAPTCHA type	
	Text-recognition	Image-recognition
Verbals	42	23
Imagers	24	19
Intermediates	14	9
Total	80	51

4.2. Task performance related to cognitive styles (User Study A)

Task performance was measured as task completion efficiency and effectiveness. For task completion efficiency, two separate analyses were performed: (i) comparison of solving times between all CAPTCHA sessions that also included more than one attempts to solve the challenge; and (ii) comparison of solving times between CAPTCHA sessions that were solved at first attempt. Complementary data measures such as number of CAPTCHA refreshes are also reported.

4.2.1. Task completion efficiency of all CAPTCHA sessions

A three by two way factorial analysis of variance (ANOVA) was conducted aiming to examine main effects and interactions between the users' cognitive styles (i.e., Verbal, Imager and Intermediate) and CAPTCHA preference (i.e., text- vs. image-recognition) over the time needed to solve a CAPTCHA challenge. We illustrate the results in Fig. 6.

As assessed by inspection of boxplots, there were five outliers in the data that were caused by sessions that included more than four attempts to solve the challenge (increasing thus significantly the total time to solve the challenge), and were removed from the current analysis. The analysis revealed that, the main effect of users' cognitive styles on time needed to solve a CAPTCHA challenge is not significant ($F(2, 125)=2.345$, $p=0.1$, *partial* $\eta^2=0.038$). In contrast, a significant main effect of the CAPTCHA challenge type (i.e., text- vs. image-recognition) with regards to the time needed to solve a challenge has been identified ($F(1, 125)=34.402$, $p<0.001$, *partial* $\eta^2=0.224$), as users solved text-recognition CAPTCHA significantly more efficient than image-recognition CAPTCHA (Fig. 6).

Furthermore, a pairwise comparison between CAPTCHA types for each cognitive styles' group (Table 5) revealed that users of the Verbal and Intermediate class performed significantly

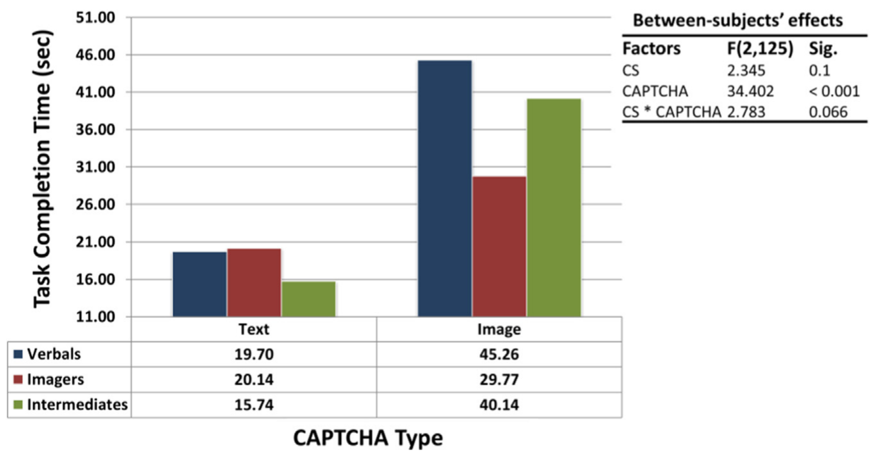


Fig. 6. Means of task efficiency per cognitive styles' group (CS) and CAPTCHA preference for all sessions.

Table 5
Pairwise comparisons of CAPTCHA types per cognitive styles' group regarding task efficiency.

Cognitive styles	(I) CAPTCHA	(J) CAPTCHA	Mean Diff. (I–J)	Sig.
Verbals	Text	Image	−25.560	< 0.001
Imagers	Text	Image	−9.629	=0.077
Intermediates	Text	Image	−24.400	=0.001

faster in text- than in image-recognition CAPTCHA challenges (Verbals: $MD = -25.560$, $SE = 4.528$; $F(1, 119) = 31.866$, $p < 0.001$, Intermediates: $MD = -24.400$, $SE = 7.323$; $F(1, 119) = 11.103$, $p = 0.001$). However, users belonging to the Imager class had no significant effect on task efficiency between text- and image-recognition CAPTCHA challenges ($MD = -9.629$, $SE = 5.394$; $F(1, 119) = 3.187$, $p = 0.077$), as they performed faster in the image-recognition CAPTCHA challenge compared to the other two groups. An interpretation of this result can be based on the fact that all users were more familiar and experienced interacting with text-recognition CAPTCHA, hence the majority of users (mostly Verbals and Intermediates) were more efficient at solving the text-recognition challenge. On the other hand, since the familiarity factor did not affect the image-recognition CAPTCHA, we have observed that the visual approach of processing and organizing information of the Imagers has positively affected their task completion efficiency compared to the Verbals. This is further supported based on a pairwise comparison between cognitive styles' groups, revealing that Imagers were significantly faster at solving image-recognition CAPTCHA compared to Verbals ($MD = -15.489$, $SE = 5.394$; $F(1, 119) = 4.158$, $p = 0.018$).

4.2.2. *Task completion efficiency of CAPTCHA sessions solved at first attempt*

The same analysis was conducted as the previous one, for cases that CAPTCHA sessions were solved at first attempt without any errors. Main aim was to analyze the users' actual cognitive processing time required to solve the challenge, since a failed attempt or refresh loads a new challenge, requiring the users to restart the cognitive process. Fig. 7 illustrates the means of task efficiency per cognitive styles' group and preference towards CAPTCHA.

The new analysis revealed that the main effect of users' cognitive styles on time needed to solve a CAPTCHA challenge is significant ($F(2, 102) = 5.276$, $p = 0.007$, *partial* $\eta^2 = 0.099$). Similar to the previous analysis, the effect of the CAPTCHA challenge type (i.e., text- vs. image-recognition) on the time needed to solve a CAPTCHA challenge was significant ($F(1, 102) = 27.450$, $p < 0.001$, *partial* $\eta^2 = 0.222$), as users solved the text-recognition challenge more efficiently than the image-recognition challenge. Furthermore, there was an interaction effect between users' cognitive styles and CAPTCHA type on the time needed to solve the challenge ($F(1, 102) = 5.358$, $p = 0.006$, *partial* $\eta^2 = 0.1$).

To this end, the results provide initial indications that cognitive styles play an important role on CAPTCHA solving time and that image-recognition CAPTCHA could be provided as an alternative CAPTCHA mechanism to Imager users since no significant differences were observed with the text-recognition CAPTCHA (which had an additional advantage of the familiarity factor since users were more experienced with text-recognition challenges). In addition, Imagers were significantly faster than Verbals in solving image-recognition challenges.

4.2.3. *Task completion effectiveness*

Task completion effectiveness was measured as the success rate of the CAPTCHA session; for example, when a user solved the CAPTCHA at first attempt, the success rate value is 100%, whereas for a user that solved the challenge at third attempt, the success rate value is 33%. A three by two way factorial analysis of variance (ANOVA) was conducted using cognitive styles (Verbal/Imager/Intermediate) and CAPTCHA preference (text and image) as independent variables and CAPTCHA success rate as the dependent variable. Fig. 8 illustrates the success rate per cognitive styles' group and CAPTCHA type. Table 6 also summarizes the total number of attempts per cognitive styles' group and CAPTCHA type.

The analysis revealed that, the main effect of users' cognitive styles on success rate to solve a CAPTCHA challenge is not significant ($F(2, 131) = 0.796$, $p = 0.374$, *partial* $\eta^2 = 0.006$). In addition, there was no main effect of the CAPTCHA challenge type (i.e., text- vs. image-recognition) on the CAPTCHA success rate ($F(1, 131) = 2.143$, $p = 0.122$, *partial* $\eta^2 = 0.033$).

The results might be explained by the fact that the majority of sessions were successfully completed at first attempt. Nevertheless, based on the descriptive statistics, we observe that Verbals and Intermediates have better success rates in the case of text-recognition challenges compared to the image-recognition. In the case of Imagers, minimal differences in success rate exist between the two CAPTCHA types. Also worth mentioning is the fact that Intermediates had the highest success rates in both CAPTCHA types compared to the other user groups, with a 100% success rate of all Intermediate users in the case of text-recognition CAPTCHA.

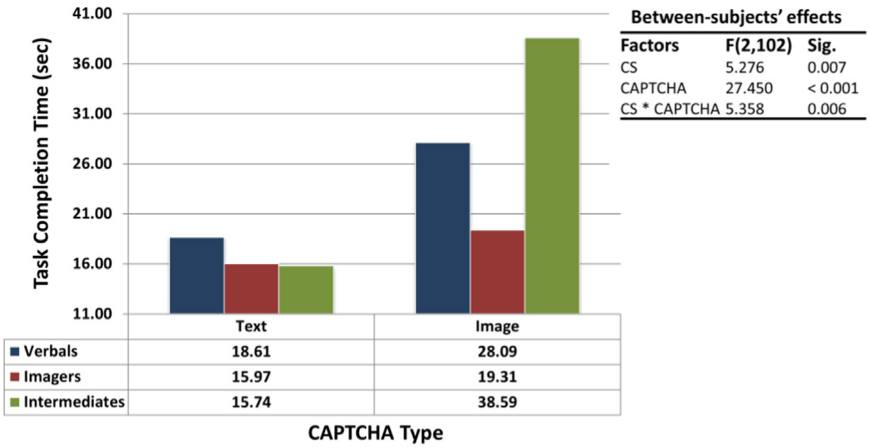


Fig. 7. Means of task efficiency per cognitive styles' group (CS) and CAPTCHA preference for sessions solved at first attempt.

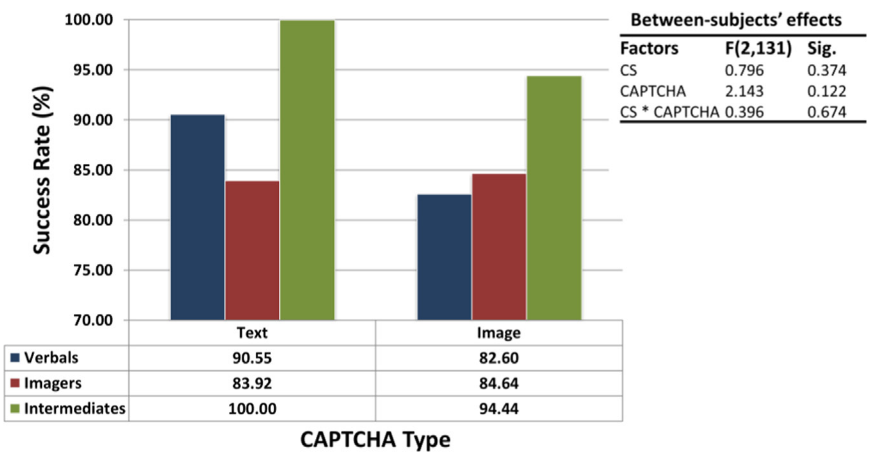


Fig. 8. Means of success rate per cognitive styles' group (CS) and CAPTCHA preference.

Table 6
Number of attempts per cognitive styles' group and CAPTCHA type.

Cognitive styles	Attempts	CAPTCHA type		
		Text-recognition	Image-recognition	Total
Verbals	1	36	16	52
	2	2	4	6
	3	1	3	4
	> 4	3	0	3
Imagers	1	18	14	32
	2	2	3	5
	3	3	1	4
	> 4	1	1	2
Intermediates	1	14	8	22
	2	0	1	1

Table 7
Number of refreshes per cognitive styles' group and CAPTCHA type.

	Verbals	Imagers	Intermediates	Total
Text	1	0	6	7
Image	3	0	0	3

4.2.4. Complementary data measures

The number of refreshes was recorded as complementary data measures. Table 7 summarizes the number of refreshes per cognitive styles' group and CAPTCHA type.

Imagers did not refresh any challenge in both CAPTCHA types, whereas 3 sessions of Verbals (1 user with 2 refreshes, 1 user with 1 refresh) initiated a refresh in an image-recognition CAPTCHA, while 1 refresh was initiated in a text-recognition CAPTCHA. In the case of Intermediates, 6 refreshes were recorded in the case of text-recognition, among them, 1 user initiated 2 consecutive refreshes in the same session. The Kruskal–Wallis H and the Mann–Whitney *U* test respectively did not reveal significant differences between the 3 cognitive styles' groups and the 2 CAPTCHA types on the number of refreshes.

4.3. Task performance related to cognitive processing abilities (User Study B)

In the second user study we aimed to investigate whether cognitive processing abilities of users affect task efficiency and effectiveness of both text- and image-recognition CAPTCHA. Two separate analyses were performed for text-recognition and image-recognition CAPTCHA interactions since the allocation of CAPTCHA type (text or image) was based on the Verbal/Imager cognitive styles' dimension. Given the between-subject study design (the allocation of CAPTCHA complexity was split randomly to users based on their cognitive processing abilities), we conducted the analysis of variance (ANOVA) test. Both task efficiency analyses included times from all CAPTCHA sessions (including sessions that needed more than one attempt to complete) since the same trend was observed for CAPTCHA sessions that were completed at first attempt and those

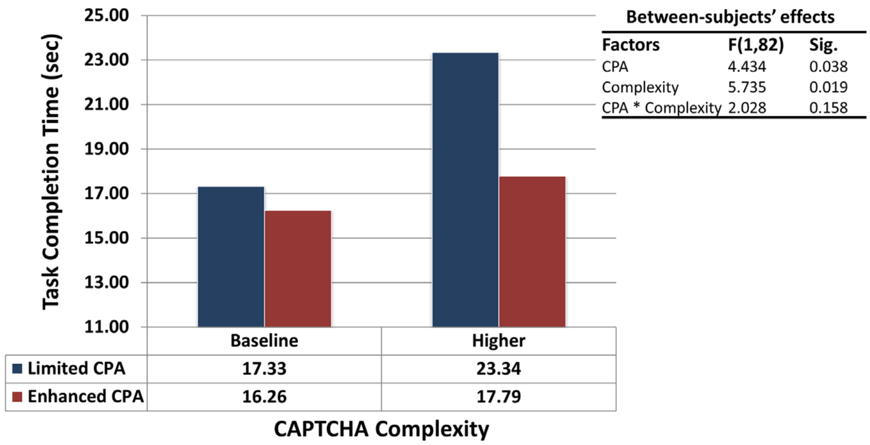


Fig. 9. Users' cognitive processing abilities (CPA) and text-recognition CAPTCHA task efficiency.

Table 8
Pairwise comparisons of CAPTCHA complexity per cognitive processing abilities' group regarding task efficiency.

Cognitive processing abilities	(I) CAPTCHA	(J) CAPTCHA	Mean Diff. (I–J)	Sig.
Limited	Baseline	Higher	−6.006	=0.013
Enhanced	Baseline	Higher	−1.526	=0.466

that needed more than one attempt. Additional data measures such as number of CAPTCHA refreshes are also reported.

4.3.1. *Task completion efficiency and effectiveness of text-recognition CAPTCHA*

A two by two way factorial analysis of variance (ANOVA) was conducted aiming to examine main effects and interactions between the users' cognitive processing abilities (i.e., limited, enhanced) and CAPTCHA complexity level (i.e., baseline, higher) over the time needed to solve a text-recognition CAPTCHA challenge. We summarize the results in Fig. 9.

Results indicate that complexity level (baseline/higher) has a main effect on task completion time ($F(1, 82)=5.735$, $p=0.019$, *partial* $\eta^2=0.068$). Such a result was rather expected given the increased number and added complexity of the characters. In addition, there was a main effect of cognitive processing abilities on time to complete ($F(1, 82)=4.434$, $p=0.038$, *partial* $\eta^2=0.054$) since users with enhanced cognitive processing abilities were significantly faster in solving both types of CAPTCHA complexity designs compared to users with limited cognitive processing abilities. Furthermore, there was no interaction effect between cognitive processing abilities and complexity level on the time to complete ($F(1, 82)=2.028$, $p=0.158$, *partial* $\eta^2=0.025$). Pairwise comparisons between baseline and high complexity levels (Table 8) revealed that users with limited cognitive processing abilities completed the baseline level CAPTCHA significantly faster compared to the high complex CAPTCHA ($MD=-6.006$, $SE=2.355$; $F(1, 78)=6.504$, $p=0.013$, *partial* $\eta^2=0.077$). In contrast, no significant differences were observed between baseline and higher levels of CAPTCHA complexity for users with enhanced cognitive processing abilities ($MD=-1.526$, $SE=2.085$; $F(1, 78)=0.536$, $p=0.466$, *partial* $\eta^2=0.007$).

Such a result suggests that CAPTCHA with higher complexity could be provided to users with enhanced cognitive processing abilities given that the usability in terms of task efficiency is not significantly decreased. In the context of a personalization system, increasing the CAPTCHA complexity to users with enhanced cognitive processing abilities could increase the security level of the CAPTCHA at a rather small cost to usability. In contrast, it is suggested not to provide highly complex CAPTCHA (but rather a baseline complexity level) to users with limited cognitive processing abilities since this would increase significantly the CAPTCHA task completion time.

For task effectiveness, a two by two way factorial analysis of variance (ANOVA) was conducted to investigate whether cognitive processing abilities of users have an effect on success rate (Fig. 10). Results revealed a main effect of complexity on success rate ($F(1, 82)=4.050$, $p=0.048$, *partial* $\eta^2=0.049$) since the success rate of baseline complexity CAPTCHA was overall higher compared to the higher complex CAPTCHA. There was also a main effect of cognitive processing abilities on the success rate of text-recognition CAPTCHA ($F(1, 82)=4.109$, $p=0.046$, *partial* $\eta^2=0.05$) since users with enhanced cognitive processing abilities had a higher success rate in both types of complexity levels.

Results suggest that highly complex text-recognition CAPTCHA hinder the usability in terms of task effectiveness for users with limited cognitive processing abilities since they were significantly less effective in solving the highly complex CAPTCHA compared to the baseline complexity CAPTCHA (Baseline-Higher CAPTCHA: $MD=19.908$, $SE=8.580$; $F(1, 78)=5.383$, $p=0.023$, *partial* $\eta^2=0.065$). In addition, pairwise comparisons between limited and enhanced user groups revealed that in the case of baseline level CAPTCHA, no significant differences were observed between the two user groups (Limited-Enhanced: $MD=-3.241$, $SE=8.026$; $F(1, 78)=0.163$, $p=0.687$, *partial* $\eta^2=0.002$), whereas in the case of high level CAPTCHA, users with enhanced cognitive processing abilities had significant higher success rates than those with limited abilities (Limited-Enhanced: $MD=-19.992$, $SE=8.181$; $F(1, 78)=5.972$, $p=0.017$, *partial* $\eta^2=0.071$).

4.3.2. *Task completion efficiency and effectiveness of image-recognition CAPTCHA*

The same ANOVA analysis was conducted for image-based interactions as the one conducted for text-based interactions. We summarize the results in Fig. 11.

Similarly to the text-recognition CAPTCHA analysis, complexity level (baseline/higher) has a main effect on task completion time ($F(1, 43)=4.730$, $p=0.036$, *partial* $\eta^2=0.108$) since users across all groups performed faster in the baseline complexity image-recognition CAPTCHA. Furthermore, cognitive processing abilities did not have a main effect on task completion efficiency ($F(1, 43)=0.856$, $p=0.361$, *partial* $\eta^2=0.021$). Also, there was no interaction

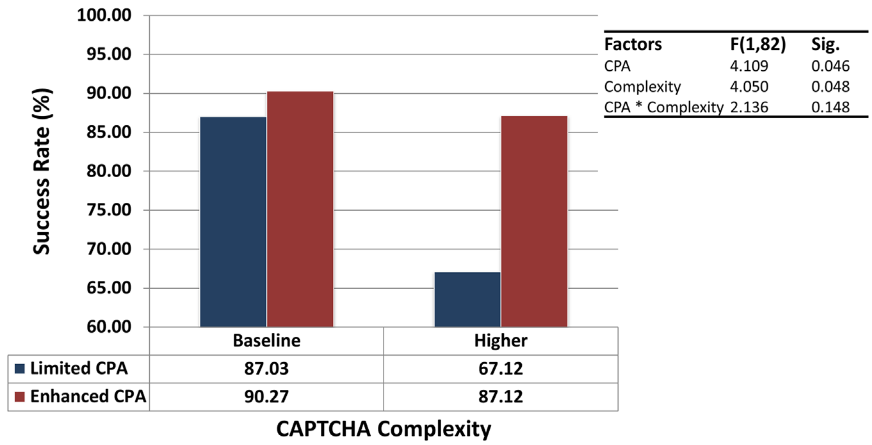


Fig. 10. Users' cognitive processing abilities (CPA) and text-recognition CAPTCHA task success rate.

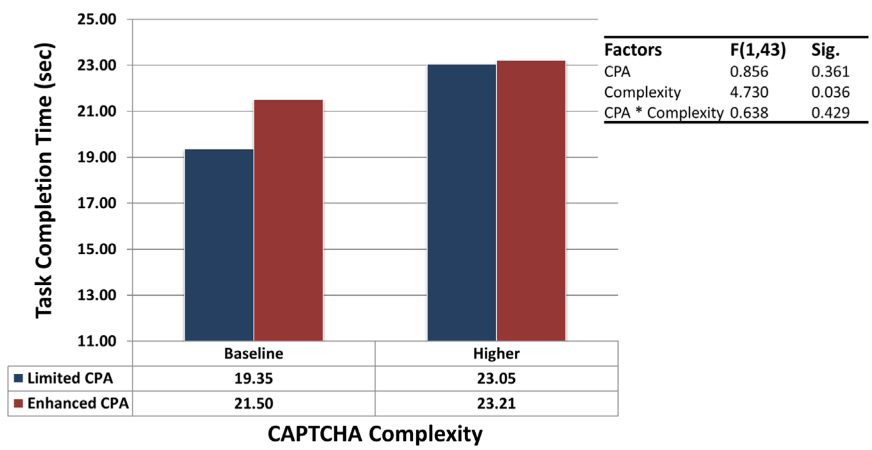


Fig. 11. Users' cognitive processing abilities (CPA) and image-recognition CAPTCHA task efficiency.

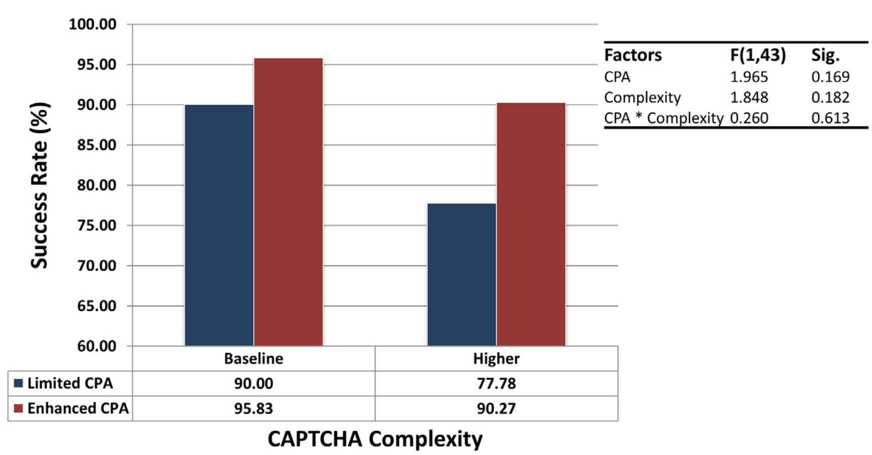


Fig. 12. Users' cognitive processing abilities (CPA) and image-recognition CAPTCHA task success rate.

effect between cognitive processing abilities and CAPTCHA complexity on time to complete the image challenge ($F(1, 43)=0.638$, $p=0.429$, *partial* $\eta^2=0.016$). Based on the results we suggest providing the same baseline ASIRRA version (12 colored images) to both users with limited and enhanced cognitive processing abilities. More images and additional security measures could be provided for increasing the security of ASIRRA that would however decrease considerably its task completion efficiency for both user types.

Finally, regarding task effectiveness (Fig. 12), based on descriptive statistics we observe that the success rate is decreased from baseline to higher complexity image-based CAPTCHA (especially for users with limited cognitive processing abilities). However, based on the ANOVA test analyses, results did not reveal a main effect of complexity on success rate ($F(1, 43)=1.848$, $p=0.182$, *partial* $\eta^2=0.045$). Also, there was no main effect of cognitive processing abilities on the success rate of image-recognition CAPTCHA ($F(1, 43)=1.965$, $p=0.169$, *partial* $\eta^2=0.048$).

Table 9
Number of refreshes per cognitive processing abilities' (CPA) group (limited/enhanced), CAPTCHA type (text/image) and complexity level (baseline/higher).

CAPTCHA (complexity)	Limited CPA	Enhanced CPA	Total
Text (baseline)	1	2	3
Text (higher)	3	4	7
Image (baseline)	2	2	4
Image (higher)	1	0	1

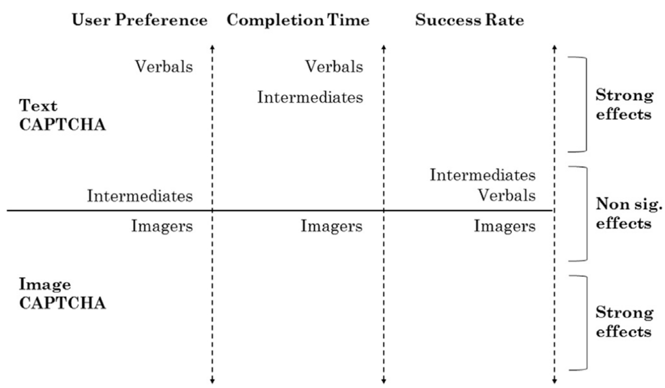


Fig. 13. Main effects of users' cognitive styles on CAPTCHA preference and performance (User Study A).

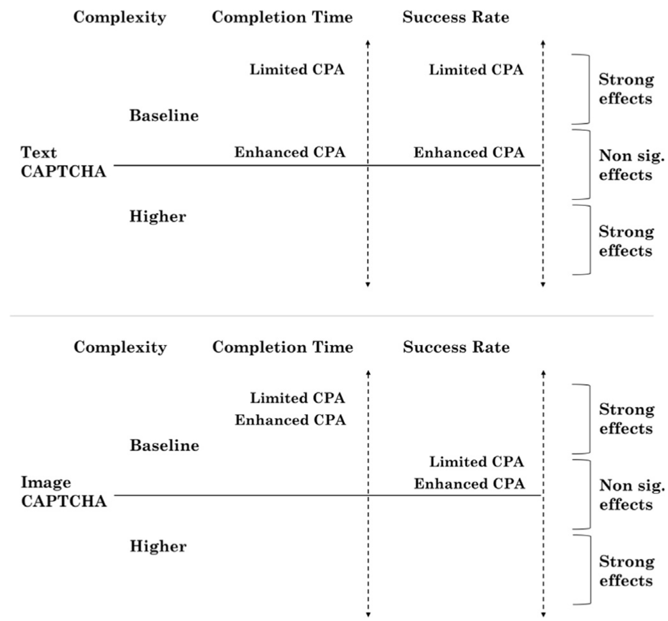


Fig. 14. Main effects of users' cognitive processing abilities (CPA) on CAPTCHA performance (User Study B).

4.3.3. Complementary data measures

The number of refreshes was recorded as complementary data measures. Table 9 summarizes the number of refreshes per user group, CAPTCHA type and complexity level. The Mann–Whitney *U* test did not reveal significant differences between the user groups and CAPTCHA complexity levels on the number of refreshes.

5. Discussion of results

Analysis of results demonstrated several main effects of cognitive processing characteristics of users on preference and task performance of different visual designs of CAPTCHA challenges. In this section, we summarize the results of the studies and accordingly propose guidelines that service providers could take into consideration for designing and developing personalized CAPTCHA challenges with the aim to improve the usability and the overall user experience during such interactions. Figs. 13 and 14 respectively provide a visual illustration of the main effects of users' cognitive styles (Verbal/Imager/Intermediate) on CAPTCHA preference and performance in User Study A, and the main effects of users' cognitive processing abilities (limited/enhanced) on CAPTCHA performance in User Study B. In both figures, the vertical lines show the effect's strength of the particular cognitive characteristic (i.e., Verbal/Imager/Intermediate, and limited/enhanced cognitive processing abilities (CPA)) on the usability metrics investigated (i.e., user preference, task completion, success rate) with respect to the different CAPTCHA designs (i.e., text vs. image, and baseline vs. higher complexity). The horizontal dark line distinguishes the CAPTCHA designs (text vs. image, baseline vs. higher complexity).

Regarding user preference (User Study A), participants in general preferred significantly text- than image-recognition CAPTCHA challenges. This finding can be explained by taking into consideration that the majority of Web application providers utilize text-recognition CAPTCHA (Bursztein et al., 2010), and thus, users are more familiar in solving text- than image-recognition CAPTCHA challenges. Results also revealed a significant preference for users belonging to the Verbal class to choose text- than image-recognition challenges and, as the sample increases there is a growing trend for users belonging to the Imager class to prefer image-recognition challenges. Results of task efficiency (User Study A) revealed that Verbals and Intermediates were significantly more efficient when interacting with text-recognition CAPTCHA, whereas in the case of Imagers, no significant differences in performance were observed between the two variations of CAPTCHA since they solved image-recognition CAPTCHA challenges much faster than the other two cognitive style groups. This result indicates that Imagers were positively affected by their cognitive style of processing more efficiently graphical than text-based information, underpinning that cognitive styles are important to be considered in the design of CAPTCHA challenges. Accordingly, we suggest that image-recognition CAPTCHA challenges could be a viable alternative to current text-recognition challenges for Imager users. On the other hand, it is suggested to provide text-recognition CAPTCHA challenges to Verbals and Intermediates.

With regards to task performance (efficiency and effectiveness) in User Study B, this research work suggests providing a baseline complexity level of text-based CAPTCHA to users with limited cognitive processing abilities since they were negatively affected by the increase of character distortion and noise in the challenge. In contrast, a highly complex text-based CAPTCHA could be provided to users with enhanced cognitive processing abilities in order to increase CAPTCHA security at a rather non-significant negative cost to usability.

Regarding image-recognition CAPTCHA, results suggest that service providers should bear in mind that increasing the complexity level of image-recognition CAPTCHA (e.g., increase the number of images), affects negatively the task completion efficiency and effectiveness for both types of users (limited/enhanced cognitive processing abilities). A baseline 12-image colored CAPTCHA challenge should be provided as proposed in Elson et al. (2007) to all types of users since task completion is significantly more efficient and effective compared to the 14-image greyscale challenge.

Finally, although Imagers had significant improved task completion efficiency in image-recognition CAPTCHA compared to

Verbals and Intermediates, results have shown that in general, users were solving the image-recognition CAPTCHA challenge (Microsoft ASIRRA) less efficiently compared to the text-recognition CAPTCHA. Based on our observations and experiences while conducting the studies, this might be caused due to the interaction design of the Microsoft ASIRRA implementation. In particular, the current version of Microsoft ASIRRA illustrates 12 small (40 pixel × 40 pixel) colored images and users are required to hover the mouse pointer on each image in order to view the larger version of the image, which is rather time consuming. Accordingly, we suggest not using the mouse hover technique but instead illustrating larger (120 pixel × 120 pixel) and responsive images by leveraging current CSS3 features (World Wide Web Consortium (W3C), 2014) for adapting the size of the images according to the device's screen size dimensions (mobile touch-based vs. monitor of desktop computer). Accordingly, based on the usability evaluation study of ASIRRA (Elson et al., 2007), the typical response time for recognizing a 14,400 pixel image (120 pixel × 120 pixel) is estimated to be 10.2 s for solving a 12-image CAPTCHA challenge, with a 98.5% per image accuracy and an overall success rate of 83.4%. Given the reported main effects of cognitive styles (Verbal/Imager/Intermediate) on task efficiency of text- and image-recognition CAPTCHA, in addition with the suggested visual and interaction design enhancements of the ASIRRA challenge, we expect that personalized CAPTCHA types (e.g., image-recognition for Imagers) would significantly improve task completion efficiency.

6. Implications

The discussion of results suggests that human cognitive factors and CAPTCHA design factors affect task completion performance and user preference of CAPTCHA mechanisms. Individuals with certain cognitive styles and cognitive processing abilities have particular characteristics that influence their performance when interacting with different CAPTCHA types and visual design complexity. This denotes that CAPTCHA mechanisms could embrace both text-based and image-based challenges, bootstrapped on the unique cognitive characteristics and abilities of users, with the aim to improve the task usability and eventually provide a positive user experience.

From this perspective, there is a need to transform the interdependencies among human and CAPTCHA design factors into formal representations for modeling the users' individual characteristics and accordingly provide adaptive and personalized CAPTCHA tasks. Recently, we proposed a generic formalization framework in Fidas et al. (2015) and a human cognitive factor-based formalization framework in Belk et al. (2015) which address an optimization problem related to assigning the most optimized CAPTCHA mechanism, given specific human and technology factors, and CAPTCHA design properties and attributes (e.g., an individual being Verbal or Imager). The formalization of the generic framework is expressed as follows:

> Let U denote a set of users $U=\{u_1,u_2,\ldots,u_n\}$. Let FC denote a set of factors which are maintained by the service provider $FC=\{hfc_1,hfc_2,\ldots,hfc_n,\ tfc_1,tfc_2,\ldots,tfc_n,\ cdfc_1,cdfc_2,\ldots,cdfc_n\}$. Let $UCM_j(u_i)$ denote a set of factors of the individual context model of user u_i. The result of $UCM_j(u_i)$ is a set of triplets of the form (u_i, fc_i, val), where j is the triplet identifier, u_i is the user, fc_i is the factor of the model and val is the value of the factor fc_i, where val can be any value type (e.g., Numeric, String, Boolean, etc.). Let CR denote a set of context rules which are maintained by the service provider $CR=\{cr_1,cr_2,\ldots,cr_n\}$. Each context rule is based on a decision making model which has one hypothesis part related to human and technology factors and precisely one decision part related to CAPTCHA design factors. As such, the service administrator could select certain factor properties/attributes, set the desired values and relate them with the appropriate Boolean logical connectives ($Blc=\{and,or,not,xor,\ldots\}$) and Operators ($Opr=\{=,==,<,>,!=,\ldots\}$) in order to construct fully parenthesized expressions of arbitrary complexity that can be applied to a group of users or to specific individuals following deductive or inductive reasoning approaches (Fidas et al., 2015).

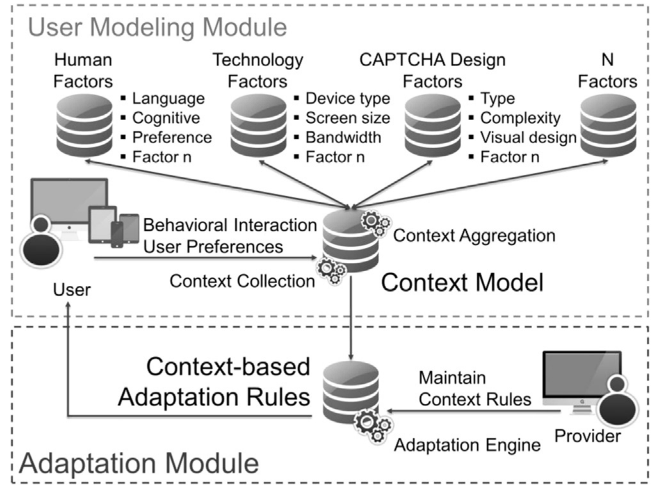


Fig. 15. Conceptual design of an extensible framework specializing on delivering personalized CAPTCHA challenges (Fidas et al., 2015).

In the context of this research, the aforementioned formalization could be realized and extended through the following main factors: *human factors*: cognitive styles and cognitive processing abilities; and *design factors*: CAPTCHA type, CAPTCHA visual complexity. Thus, the set of factors *FC* is extended as $FC=\{cs, cpa, type, complexity\}$, where *cs* indicates the user's cognitive styles ($val=\{Verbal|Imager|Intermediate\}$), *cpa* indicates the user's cognitive processing abilities ($val=\{limited|enhanced\}$), *type* indicates the recommended CAPTCHA type ($val=\{text|image\}$), and *complexity* indicates the recommended CAPTCHA visual complexity ($val=\{baseline|higher\}$).

The design of the framework conceptually consists of two main modules; the *user modeling module* that is responsible to elicit and store the user's overall context of use during interaction (human and technology specific), and the *adaptation module* that is responsible to map human factors with CAPTCHA design factors aiming to deliver the most optimized CAPTCHA challenge to each user. Fig. 15 illustrates the conceptual design of the framework.

6.1. Recommendation rules

The main results of this study could be transformed into specific context-based recommendation rules (*CR*), and further applied in a procedure for recommending a particular CAPTCHA type and visual complexity by considering the users' cognitive processing styles and abilities. Algorithm 1 presents the procedure in pseudo-code for recommending a CAPTCHA type and complexity given a user's individual context model ($UCM_j(u_i)$).

Algorithm 1. CAPTCHA Recommendation (*CAR*).

Input: A set of individual context models *UCM* and a user u_i.
Output: Assign CAPTCHA *type* and CAPTCHA *complexity* of user u_i with a recommendation *rt* and *rc*.
1: **procedure: CAR(*UCM*, u_i)**

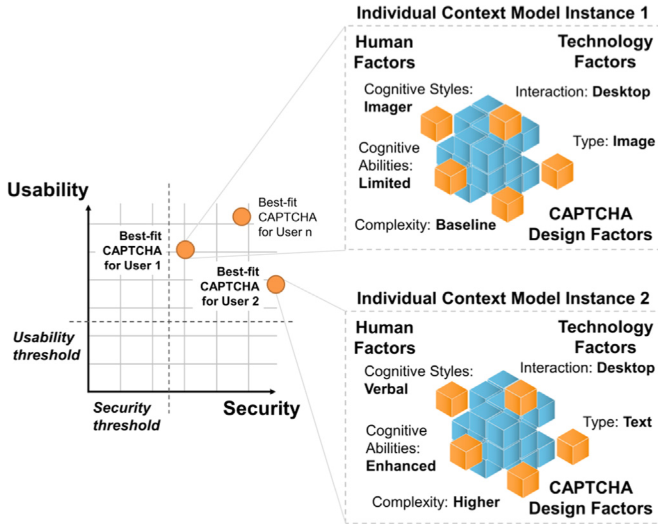


Fig. 16. Best-fit CAPTCHA challenge generation based on human factors.

```
rt=null; rc=null;
// create UF by extracting a set of tuples (fc_i, val) from UCM for user u_i
UF={ (fc_i, val): (u_k, fc_z, val) ∈ UCM and u_i=u_k }
if ( ((cs, Verbal)∩UF≠∅ or (cs, Intermediate)∩UF≠∅) and
        (cpa, limited)∩UF≠∅ )
    rt=textual; rc=baseline;
else if ( ((cs, Verbal)∩UF≠∅ or (cs, Intermediate)∩UF≠∅) and
        (cpa, enhanced)∩UF≠∅ )
    rt=textual; rc=higher;
else if ( (cs, Imager)∩UF≠∅ and
        ((cpa, limited)∩UF≠∅ or (cpa, enhanced)∩UF≠∅) )
    rt=image; rc=baseline;
UCM=UCM∪(u_i, type, rt);
UCM=UCM∪(u_i, complexity, rc);
end procedure
```

The procedure runs during a CAPTCHA session of user u_i. The procedure extracts the values of all the factors (*UF*) from the user's individual context model and accordingly recommends a specific CAPTCHA type and complexity based on several context-based rules which reflect the observed main effects of this user study. For example, according to Algorithm 1, when users have a Verbal cognitive style, with limited cognitive processing abilities, the procedure provides a text-recognition CAPTCHA with a baseline complexity level of character distortion and noise. This recommendation rule was based on the reported user studies' results that have shown that Verbal users significantly prefer and solve faster text-recognition challenges, compared to image-recognition challenges. In addition, users of the Verbal class having limited cognitive processing abilities are provided with a baseline complexity visual design CAPTCHA since results have shown that the time to solve the challenge significantly increases, and the task completion effectiveness significantly decreases as the visual design becomes more complex. On the other hand, Verbal users that have enhanced cognitive processing abilities are provided with a higher complex visual design since results have shown that these users do not perform significantly different between the baseline and higher complex design. Furthermore, when users have an Imager cognitive style, regardless of their cognitive processing abilities, an image-recognition CAPTCHA is recommended with baseline complexity since results of the studies have shown that Imager users interact equally well on both text- and image-recognition CAPTCHA, regardless of their familiarity with text-recognition CAPTCHA, and because there is a growing trend of Imager users preferring image-recognition CAPTCHA. Fig. 16 illustrates instances of different user models and their respective recommendation based on the procedure.

Accordingly, different CAPTCHA challenges are provided to the users based on the respective factors of each individual context model as well as based on custom requirements of the service provider and the application domain. For example, User 1 in Fig. 16 will receive an image-recognition CAPTCHA based on his cognitive styles (Imager), and a baseline complexity CAPTCHA design due to his limited cognitive processing abilities. Accordingly, based on the recommendation, the usability of the task is increased for User 1, and the security is not compromised based on a custom security threshold defined by the service provider. On the other hand, User 2 will receive a text-recognition CAPTCHA (Verbal) with a higher complex CAPTCHA design due to his increased cognitive processing abilities. This way the security is increased and the usability is not compromised according to a minimum usability threshold defined by the service provider.

The aforementioned recommendation rules could be further extended with other human, technology and design factors. For example, the recommendation rules could incorporate the device factor since a recent study of Wismer et al. (2012) revealed the users' positive attitude and preference towards solving image-recognition CAPTCHA on mobile devices, in contrast to text- and speech-recognition CAPTCHA. Thus, the suggested recommendation rules might change for users interacting with mobile devices (e.g., in case the user is Verbal and interacts on a smartphone) and eventually recommend an image-recognition CAPTCHA challenge instead of a text-recognition CAPTCHA. On the other hand, based on another recent study of Reynaga et al. (2015) that explored the usability of various CAPTCHA types on smartphones, participants with familiarity and experience in interacting with smartphone virtual keyboards, tend to prefer simple and quick schemes such as NuCAPTCHA and Emerging rather than image-recognition CAPTCHA. In this context, the recommendation rules should be further extended with the users' familiarity and experience in regards with smartphone virtual keyboard interactions.

In this context, other factors such as type of challenge (e.g., animated text) (design factor), user's familiarity and experience (human factor) and screen size (technology factor) should be investigated in future studies and incorporated as factors in the user model in order to more inclusively represent the users' characteristics and their context of interaction and thus make more precise CAPTCHA challenge recommendations.

7. Validity and limitations

This research work entails a number of limitations that are inherent to the multi-dimensional character and complexity of the factors investigated. Aiming to increase the validity of this research work, we addressed its internal, ecological, and external validity. Internal validity reflects the accuracy of data and the conclusions drawn based on this data, ecological validity requires that the experimental design, procedure and setting of the study must approximate the real-life context that is under investigation (Brewer, 2000), and external validity indicates whether the data and the conclusions drawn can be generalized to a wider extend (Cook and Campbell, 1979).

An important limitation is related to the rather limited number and non-varying user profiles of the sample since undergraduate students were recruited for conducting the studies. Furthermore, the sample included users who were more familiar with text-recognition CAPTCHA than with image-recognition CAPTCHA. On the other hand, with the aim to increase internal validity we recruited a sample of participants with similar profiles and experienced, rather than novice users that would lack familiarization with CAPTCHA mechanisms. Given the highly complex nature of cognitive factors and the study, the main aim of this approach was to control any other contextual factors (e.g., experience, educational background, age, nationality) that might influence the study procedure and the results, and thus isolate and investigate the impact of specific human cognitive factors on CAPTCHA design factors. Nonetheless, future work entails extending the investigation with additional factors aiming to increase our understanding about the effects of cognitive factors on CAPTCHA and their possible interaction with other contextual factors (e.g., age, culture, device, etc.). In particular, bearing in mind that prior research has shown cross-cultural differences in cognitive styles (Western vs. Eastern societies (Cui et al., 2013; Varnum et al., 2010), African American vs. South African (Engelbrecht et al., 1997)), future work entails investigating the effects of inter-cultural differences on CAPTCHA across different countries and continents.

Furthermore, there has been an effort to increase ecological validity of the research since the CAPTCHA challenges were integrated in a real-world setting and the participants were involved at their own physical environments without the intervention of any experimental equipment or person. In particular, participants were required to solve the CAPTCHA challenges as a secondary task during the real-life primary task of interacting with the course's Web-site.

Two particular CAPTCHA mechanisms were investigated although numerous other mechanisms and features exist in practice and in the academic literature. Current CAPTCHA mechanisms entail a number of features that need further investigation, such as the types of text-based challenges (e.g., static vs. animated text in NuCAPTCHA, 2015), the interaction design of the CAPTCHA mechanism (e.g., drag-and-drop action in SweetCAPTCHA, 2015), alternative cognitive tasks (e.g., determine the upright orientation of an image in Google's What's Up CAPTCHA (Gossweiler et al., 2009)). These features affect differently both the usability and security aspects of the CAPTCHA mechanism and future studies are necessary to increase the validity of this research.

Nevertheless, in order to address the external validity of this research and generalize the conclusions drawn to a wider extent, we have carefully chosen the particular CAPTCHA schemes for the studies with the aim to practically cover a high number of existing CAPTCHA schemes. In particular, we have chosen the investigation of traditional text-recognition CAPTCHA mechanisms since these are currently the most popular and widely applied CAPTCHA mechanisms (Bursztein et al., 2014), and image-recognition CAPTCHA (and more specifically ASIRRA) since this belongs to a broad image-recognition CAPTCHA category ("the distinguishing CAPTCHA") which is considered among the three main image-recognition CAPTCHA scheme categories (i.e., naming images, distinguishing images, identifying anomalies) (Chew and Tygar, 2004). Thus, we believe that the reported main effects can be applied and affect a high number of alternative text- and image-recognition CAPTCHA that belong to the same CAPTCHA scheme categories. For example, given that the reported results revealed that Verbal users prefer and perform significantly faster in static text-recognition challenges (than image-based), we expect that such a result would apply as well on animated text-recognition challenges (e.g., NuCAPTCHA) since in both CAPTCHA challenges, the information is processed primarily by utilizing the verbal cognitive sub-system.

With regards to the suggested recommendation rules, we stress that these embrace new challenges from the security and technology perspective that need closer attention. Given that the suggested recommendation rules depend on the factors of the individual context model, the main skepticism on the practical feasibility of such an approach is focused on the required prior knowledge of the system on the users' cognitive characteristics, which are necessary to personalize the CAPTCHA task. In this context, implicit user data collection methods for cognitive factors' elicitation could be based on the users' interactions with the system. Such methods would increase user acceptance of the approach since the cognitive characteristics could be transparently elicited based on the users' interactions with the system, without requiring to conduct any additional psychometric tests that would add a burden to the users. Accordingly, the users' cognitive characteristics could be implicitly inferred by tracking their navigation sequence in particular sections of the system (Belk et al., 2013), by tracking their behavior with navigation tools (hierarchical maps or alphabetical index) (Chen and Liu, 2008), or based on the usage of search tools (basic or advanced search) (Chan et al., 2014).

Finally, from the security perspective, given that various CAPTCHA schemes entail different security strengths and weaknesses (Bursztein et al., 2011; Zhu et al., 2010), the recommendation of a particular CAPTCHA type and complexity level would change the security metrics of the mechanism. Accordingly, the recommendation rules could be further extended with several factors (e.g., additional IP monitoring techniques) defined by the system administrator depending on the application domain and custom requirements. In this respect, future work entails further investigating additional factors that affect security, in combination with usability factors aiming to validate the proposed human factors-based user model and the efficacy of the proposed personalization approach.

8. Conclusions and future work

The purpose of this paper is to present results of a user-centered research endeavor which investigated human cognitive differences in information processing and their effects on user preference and task performance of different CAPTCHA designs. In this context, two user studies were designed. Both of the presented studies entailed a psychometric-based survey for eliciting the users' cognitive processing characteristics, and an ecological valid interaction scenario with two complementary types of CAPTCHA (text and image). The CAPTCHA challenges embraced different levels of complexity in terms of number of characters/ images and added distortion. Results of this research provide evidence that specific human cognitive factors have a main impact on users' preference and task efficiency and effectiveness of CAPTCHA challenges.

The contribution of the paper entails two important aspects: theory and application. Regarding theory, the presented study provides evidence that socio-cognitive theories, like the reported human cognitive theories, can be considered as applicable analysis frameworks in understanding deeper CAPTCHA-related tasks. Such frameworks are necessary given the heterogeneity of users and the globalization of today's services and applications. In particular, results of the study can be interpreted under the light of cognitive processing styles and abilities as they demonstrate a main effect of human cognitive differences on task performance of different CAPTCHA designs. Verbal and Imager individuals with varying cognitive processing abilities are affected, prefer and perform differently when interacting with different CAPTCHA mechanisms and visual complexity levels. Thus, it is necessary that designers of CAPTCHA mechanisms should consider human cognitive styles and abilities of users while interacting with the system.

Regarding application, the analysis and discussion of results underpinned the necessity for versatility in the design and development of CAPTCHA mechanisms and identified several recommendation rules for delivering personalized CAPTCHA challenges driven by the observed main effects of the studies. The recommendation rules have been expressed through a dynamic multi-layered framework which allows the expression of extendable rules which consist of human and CAPTCHA design factors. In this context, future work of the authors entails to further specify, refine and evaluate the impact of recommendation rules under a functional personalization framework.

We envision that such a personalization framework would have many positive implications from the users' point of view since, providing CAPTCHA challenges, personalized to the users' cognitive styles and cognitive processing abilities would support the users' efficiency of processing information cognitively as well as decrease cognitive load, and eventually improve the user experience and user acceptance of CAPTCHA. From a security perspective, such an attempt might strengthen the security aspects of CAPTCHA since the malicious software has to pretend to be a specific individual than a general human being, which is required to work around a credible user model (Fidas et al., 2015). Overarching goal is to drive this research towards the design and development of a personalization system, specializing on recommending "best-fit" CAPTCHA challenges, with the aim to provide a viable alternative to the current state of one-size-fits-all CAPTCHA paradigm.

Acknowledgments

The work is co-funded by the PersonaWeb project under the Cyprus Research Promotion Foundation (ΤΠΕ/ΠΛΗΡΟ/0311 (ΒΙΕ)/10), the Marie Curie and EU FP7 project SocialRobot (No. 285870) and the EU FP7 project Miraculous-Life (No. 611421). We would like to thank all the students for their time participating in the studies and their valuable feedback.

References

Albert, D., Jeng, B., Tseng, C., Wang, J., 2010. A study of CAPTCHA and its application to user authentication, Proceedings of the International Conference on Computational Collective Intelligence (ICCCI 2010. Springer-Verlag, pp. 433–440.

Angeli, C., Valanides, N., Kirschner, P., 2009. Field dependence-independence and instructional-design effects on learners' performance with a computer-modeling tool. Comput. Hum. Behav. 25 (6), 1355–1366.

Baddeley, A., 1992. Working memory. Science 255 (5044), 556–559.

Baddeley, A., 2012. Working memory: theories, models, and controversies. Annu. Rev. Psychol. 63, 1–29.

Banday, M., Shah, N., 2011. Challenges of CAPTCHA in the accessibility of Indian regional websites. Proceedings of the ACM Bangalore Conference (COMPUTE 2011), 31. ACM Press, pp. 1–4.

Belk, M., Germanakos, P., Fidas, C., Andreou, P., Samaras, G., 2015. The PAC Framework: Personalized Authentication and CAPTCHA Mechanisms based on Human Cognitive Factors. Technical Report TR-15-2, Department of Computer Science, University of Cyprus.

Belk, M., Papatheocharous, E., Germanakos, P., Samaras, G., 2013. Modeling users on the world wide web based on cognitive factors, navigation behavior and clustering techniques. Syst. Softw. 86 (12), 2995–3012.

Belk, M., Germanakos, P., Fidas, C., Samaras, G., 2014. A personalisation method based on human factors for improving usability of user authentication tasks, Proceedings of the International Conference on User Modeling, Adaptation, and Personalization (UMAP 2014). Springer-Verlag, pp. 13–24.

Belk, M., Fidas, C., Germanakos, P., Samaras, G., 2012. Do cognitive styles of users affect preference and performance related to CAPTCHA challenges?, Proceedings of the ACM SIGCHI Extended Abstracts on Human Factors in Computing Systems (CHI 2012). ACM Press, pp. 1487–1492.

Bigham, J., Cavender, A., 2009. Evaluating existing audio CAPTCHAs and an interface optimized for non-visual use, Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems (CHI 2009). ACM Press, pp. 1829–1838.

Brewer, M., 2000. 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.

Brown, E., Brailsford, T., Fisher, T., Moore, A., Ashman, H., 2006. Reappraising cognitive styles in adaptive web applications, Proceedings of the International Conference of the World Wide Web (WWW 2006). ACM Press, pp. 327–335.

Bursztein, E., Bethard, S., Fabry, C., Mitchell, J., Jurafsky, D., 2010. How good are humans at solving CAPTCHAs? A large scale evaluation, Proceedings of the International Symposium on Security and Privacy. IEEE Computer Society, pp. 399–413.

Bursztein, E., Martin, M., Mitchell, J., 2011. Text-based CAPTCHA strengths and weaknesses, Proceedings of the ACM Conference on Computer and Communications Security (CCS 2011). ACM Press, pp. 125–138.

Bursztein, E., Moscicki, A., Fabry, C., Bethard, S., Mitchell, J., Jurafsky, D., 2014. Easy does it: more usable CAPTCHAs, Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems (CHI 2014). ACM Press, pp. 2637–2646.

Cegarra, J., Hoc, J., 2006. Cognitive Styles as an explanation of experts' individual differences: a case study in computer-assisted troubleshooting diagnosis. Int. J. Hum. Comput. Stud. 64 (2), 123–136.

Chan, C., Hsieh, C., Chen, S., 2014. Cognitive styles and the use of electronic journals in a mobile context. Documentation 70 (6), 997–1014.

Chellapilla, K., Larson, K., Simard, P., Czerwinski, M., 2005. Designing human friendly human interaction proofs (HIPs), Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems (CHI 2005). ACM Press, pp. 711–720.

Chen, S., Ghinea, G., Macredie, R., 2006. A cognitive approach to user perception of multimedia quality: an empirical investigation. Int. J. Hum. Comput. Stud. 64 (12), 1200–1213.

Chen, S., Liu, X., 2008. An Integrated Approach for Modeling Learning Patterns of Students in Web-Based Instruction: A Cognitive Style Perspective Article 1. ACM Trans. Comp.-Human Interact. 15 (1), 28.

Chew, M., Baird, H., 2003. BaffleText: a human interactive proof. In: Proceedings of the International Conference on Document Recognition and Retrieval (DRR 2003), pp. 305–316.

Chew, M., Tygar, J., 2004. Image recognition CAPTCHAs. In Proceedings of the International Information Security Conference (ISC 2004). Springer-Verlag, pp. 268–279.

Conway, A.R.A., Cowan, N., Bunting, M.F., 2001. The cocktail party phenomenon revisited: the importance of working memory capacity. Psychon. Bull. Rev. 8, 331–335.

Conway, A.R.A., Cowan, N., Bunting, M.F., Therriault, D.J., Minkoff, S.R., 2002. A latent variable analysis of working memory capacity, short-term memory capacity, processing speed, and general fluid intelligence. Intelligence 30, 163–183.

Cook, T., Campbell, D., 1979. Quasi-experimentation: Design and Analysis Issues for Field Settings. Houghton Mifflin Company, Boston, MA.

Cui, G., Liu, H., Yang, X., Wang, H., 2013. Culture, cognitive style and consumer response to informational vs. transformational advertising among east asians: evidence from the PRC. Asia Pac. Bus. Rev. 19 (1), 16–31.

Davidson, D., Renaud, K., Li, S., 2014. jCAPTCHA: accessible human validation, Proceedings of the International Conference on Computers Helping People with Special Needs. Springer-Verlag, pp. 129–136.

Demetriou, A., Spanoudis, G., Shayer, S., Mouyi, A., Kazi, S., Platsidou, M., 2013. Cycles in speed-working memory-g relations: towards a developmental-differential theory of the mind. Intelligence 41, 34–50.

Elson, J., Douceur, J., Howell, J., Saul, J., 2007. Asirra: a CAPTCHA that exploits interest-aligned manual image categorization, Proceedings of the International Conference on Computer and Communications Security (CCS 2007). ACM Press, pp. 366–374.

Engelbrecht, P., Engelbrecht, P., Natzel, S., Natzel, S., 1997. Cultural variations in cognitive style: field dependence vs field independence. Sch. Psychol. Int. 18 (2), 155–164.

Fidas, C., Hussmann, H., Belk, M., Samaras, G., 2015. iHIP: Towards a user centric individual human interaction proof framework, Proceedings of the ACM SIGCHI Extended Abstracts on Human Factors in Computing Systems (CHI 2015). ACM Press, pp. 2235–2240.

Fidas, C., Voyiatzis, A., 2013. On users' preference on localized vs. Latin-based CAPTCHA challenges. In: Proceedings of the IFIP TC13 International Conference on Human Computer Interaction (INTERACT 2013), 4. Springer-Verlag, pp. 358–365.

Fidas, C., Voyiatzis, A., Avouris, N., 2011. On the necessity of user-friendly CAPTCHA, Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems. ACM Press, pp. 2623–2626.

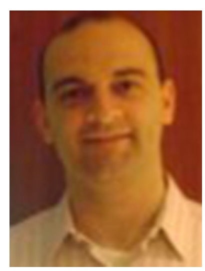
Gao, H., Liu, H., Yao, D., Liu, X., Aickelin, U., 2010. An audio CAPTCHA to distinguish humans from computers, Proceedings of the International Symposium on Electronic Commerce and Security (SECS 2010). IEEE Computer Society, pp. 265–269.

Germanakos, P., Tsianos, N., Lekkas, Z., Mourlas, C., Samaras, G., 2008. Capturing essential intrinsic user behaviour values for the design of comprehensive web-based personalized environments. Comput. Hum. Behav. 24 (4), 1434–1451.

Ghinea, G., Chen, S., 2008. Measuring quality of perception in distributed multimedia: verbalizers vs. imagers. Comput. Hum. Behav. 24 (4), 1317–1329.

Golle, P., 2008. Machine learning attacks against the Asirra CAPTCHA, Proceedings of the ACM Conference on Computer and Communications Security (CCS 2008). ACM Press, pp. 535–542.
Gossweiler, R., Kamvar, M., Baluja, S., 2009. What's up CAPTCHA?: a CAPTCHA based on image orientation, Proceedings of the International Conference on World Wide Web (WWW 2009). ACM Press, pp. 841–850.
Hale, S., Fry, A.F., 2000. Relationships among processing speed, working memory, and fluid intelligence in children. Biol. Psychol. 54, 1–34.
Holman, J., Lazar, J., Feng, J.H., D'Arcy, J., 2007. Developing usable CAPTCHAs for blind users, Proceedings of the ACM SIGACCESS Conference on Computers and Accessibility (ASSETS 2007). ACM Press, pp. 245–246.
Kluever, K., Zanibbi, R., 2009. Balancing usability and security in a video CAPTCHA, Proceedings of the ACM Symposium on Usable Privacy and Security (SOUPS 2009). ACM Press p. 11, Article 14.
Kozhevnikov, M., 2007. Cognitive styles in the context of modern psychology: toward an integrated framework of cognitive style. Psychol. Bull. 133 (3), 464–481.
Liu, Y., Ginther, D., 1999. Cognitive styles and distance education Article 5. Distance Learn. Adm. 2 (3).
MacLeod, C.M., 1991. Half a century of research on the Stroop effect: an integrative review. Psychol. Bull. 109, 163–203.
Mitchell, T., Chen, S.Y., Macredie, R., 2004. Adapting hypermedia to cognitive styles: is it necessary?, Proceedings of the Workshop on Individual Differences, International Conference on Adaptive Hypermedia and Adaptive Web-based Systems (AH 2004). Springer-Verlag.
Moradi, M., Keyvanpour, M.R., 2014. CAPTCHA and its Alternatives: A Review. Security and Communication Networks. Wiley & Sonshttp://dx.doi.org/10.1002/sec.1157.
NuCAPTCHA Inc. 2015. NuCAPTCHA—Adaptive Captcha Authentication. Retrieved on May 04, 2015 ⟨http://www.nucaptcha.com⟩.
Papanikolaou, K.A., Mabbott, A., Bull, S., Grigoriadou, M., 2006. Designing learner-controlled educational interactions based on learning/cognitive style and learner behaviour. Interact. Comput. 18 (3), 356–384.
Peterson, E., Rayner, S., Armstrong, S., 2009. Researching the psychology of cognitive style and learning style: is there really a future? Learn. Individual Differ. 19 (4), 518–523.
Polderman, T., Stins, J., Posthuma, D., Gosso, M., Verhulst, F., Boomsma, D., 2006. The phenotypic and genotypic relation between working memory speed and capacity. Intelligence 34 (6), 549–560.
Reynaga, G., Chiasson, S., van Oorschot, P. 2015. Exploring the usability of CAPTCHAS on smartphones: comparisons and recommendations. In: Proceedings of the Symposium on Network and Distributed System Security (NDSS 2015).
Rezaei, A., Katz, L., 2004. Evaluation of the reliability and validity of the cognitive styles analysis. Pers. Individual Differ. 36, 1317–1327.
Riding, R., 1991. Cognitive Styles Analysis—Research Administration. Learning and Training Technology, Birmingham, UK.
Riding, R., Cheema, I., 1991. Cognitive styles-an overview and integration. Educ. Psychol. 11 (3-4), 193–215.
Securimage v.3.5.2 2014. ⟨http://www.phpcaptcha.org⟩.
Shipstead, Z., Broadway, J., 2013. Individual differences in working memory capacity and the Stroop effect: do high spans block the words? Learn. Individual Differ. 26, 191–195.
SolveMedia 2015. SolveMedia Official Web-site. Retrieved on May 04, 2015 ⟨http://www.solvemedia.com⟩.
Stroop, J., 1935. Studies of interference in serial verbal reactions. Exp. Psychol. 18, 643–662.
SweetCAPTCHA 2015. SweetCAPTCHA—Fun and Human Friendly Captcha. Retrieved on May 04, 2015 ⟨http://sweetcaptcha.com⟩.
Toker, D., Conati, C., Steichen, B., Carenini, G., 2013. Individual user characteristics and information visualization: connecting the dots through eye tracking, Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems (CHI 2013). ACM Press, pp. 295–304.
Unsworth, N., Spillers, G.J., 2010. Working memory capacity: attention, memory, or both? A direct test of the dual-component model. Mem. Lang. 62, 392–406.
Varnum, M., Grossmann, I., Kitayama, S., Nisbett, R., 2010. The origin of cultural differences in cognition: the social orientation hypothesis. Curr. Dir. Psychol. Sci. 19 (1), 9–13.
Vikram, S., Fan, Y., Gu, G., 2011. SEMAGE: a new image-based two-factor CAPTCHA, Proceedings of the International Conference on Computer Security Applications (CCS 2011). ACM Press, pp. 237–246.
von Ahn, L., Blum, M., Langford, J., 2004. Telling humans and computers apart automatically. Commun. ACM 47, 56–60.
von Ahn, L., Maurer, B., McMillen, C., Abraham, D., Blum, M., 2008. reCAPTCHA: human-based character recognition via web security measures. Science 321 (5895), 1465–1468.
Wei, T., Jeng, A., Lee, H., 2012. GeoCAPTCHA—a novel personalized CAPTCHA using geographic concept to defend against third party human attack, Proceedings of the International Conference on Performance Computing and Communications Conference (IPCCC 2012). IEEE Computer Society, pp. 392–399.
Wismer, A., Madathil, K.C., Koikkara, R., Juang, K., Greenstein, J., 2012. Evaluating the usability of CAPTCHAs on a mobile device with voice and touch input, Proceedings of the ACM Conference on Computer and Communications Security (HFES 2012). ACM Press, pp. 1228–1232.
World Wide Web Consortium (W3C) 2014. CSS Specifications. ⟨http://www.w3.org/Style/CSS/specs⟩.
Xu, Y., Reynaga, G., Chiasson, S., Frahm, J., Monrose, F., van Oorschot, P., 2014. Security analysis and related usability of motion-based CAPTCHAs: decoding Codewords in motion. IEEE Trans. Dependable Secure Comput. 11 (5), 480–493.
Yan, J., El Ahmad, A., 2008. Usability of CAPTCHAs or usability issues in CAPTCHA design, Proceedings of the ACM Symposium on Usable Privacy and Security (SOUPS 2008). ACM Press, pp. 44–52.
Zhu, B., Yan, J., Li, Q., Yang, C., Liu, J., Xu, N., Yi, M., Cai, K., 2010. Attacks and design of image recognition CAPTCHAs, Proceedings of the ACM Conference on Computer and Communications Security (CCS 2010). ACM Press, pp. 187–200.

Marios Belk is currently a Ph.D. Candidate at the Department of Computer Science, University of Cyprus. He obtained his B.Sc. and M.Sc. in Computer Science in 2007 and 2009 respectively. His research interests lie in the areas of HCI, Adaptive Interactive Systems and Usable Security. His research work has been published in accredited scientific journals and conferences, such as Elsevier Journal of Systems and Software, Interacting with Computers, ACM CHI, UMAP, INTERACT, including also a best paper award (Springer SouthCHI 2013) and a best paper award nomination (Springer UMAP 2014).



Dr. Christos Fidas is a Lecturer at the Department of Cultural Heritage Management and New Technologies, University of Patras, Hellas. He graduated and received a PhD in Electrical & Computer Engineering in 2000 and 2004 respectively. He has registered one national patent in Information Technology. His research interests are focused in understanding users, designing with users and making technology usable. He has been involved as a user experience expert in numerous EU research and industry IT projects and has published several research articles related to IT and HCI topics.

Dr. Panagiotis Germanakos is a Research Scientist, co-founder and Deputy Head of the Semantic and Cognitive Research Adaptivity Technologies Research Group, of the University of Cyprus (2003), and Senior User Experience Researcher at SAP SE (2013). He was an ERCIM Post-doctoral Research Fellow in the Web Applications for the Future Internet Research Group in IIT-CNR (2011). He holds a PhD from the National & Kapodistrian University of Athens (2008) and his research interest is in HCI, Adaptive Cognitive Systems, Web Personalization, User Modeling, Computational Intelligence. He has over 100 publications and his work has obtained a number of awards.

Dr. George Samaras is a Professor at the Department of Computer Science, University of Cyprus. He received a Ph.D. and M.Sc. in Computer Science from Rensselaer Polytechnic Institute, USA, in 1989 and 1985 respectively. He obtained his B.Sc. degree in Mathematics at the University of Athens, Greece (1982). His research interests span the areas of Mobile and Adaptive Computing, Sensor Networks, Databases and Web Personalization Systems. He has published more than 220 articles in conferences and journals and co-authored two books, and his work was awarded with three best paper awards (ICDE 1999, AH 2008 and SouthCHI 2013).