

On Users' Preference on Localized vs. Latin-based CAPTCHA Challenges

Christos Fidas^{1,2} and Artemios G. Voyiatzis³

¹Department of Computer Science, University of Cyprus, CY-1678 Nicosia, Cyprus

²Interactive Technologies Lab, HCI Group, Electrical and Computer Engineering Department
University of Patras, GR-26504, Patras, Greece

³Industrial Systems Institute, Research Center "ATHENA"
PSP building, Stadiou Str., GR-26504, Platani Patras, Greece
christos.fidas@cs.ucy.ac.cy, bogart@isi.gr

Abstract. A Completely Automated Public Turing test to tell Computers and Humans Apart (CAPTCHA) is a widely used security mechanism for constructing a high-confidence proof that the entity interacting with a remote service is actually a human being. Stimulated by the facts that: a) nowadays CAPTCHA challenges are solely based on the Latin alphabet, b) currently Internet population consists in its majority of non-native-English speakers and c) numerous web sites consist of exclusively localized content, we conducted an empirical study aiming to examine the effect of various factors on users' preference in solving localized vs. Latin-based text CAPTCHA challenges. The study embraced a between-subject design using a self-developed localized CAPTCHA mechanism, capable of producing text challenges based on the participants' native alphabet. A total of 384 non-native English speakers participated in the frame of the reported study which followed an ecological valid experimental design. Analysis of interaction results provides interesting insights which can be taken into consideration for designing more usable CAPTCHA mechanisms.

Keywords: CAPTCHA, Usability, Security, Native Language (non-Latin) systems, localized CAPTCHA.

1 Introduction

A Completely Automated Public Turing test to tell Computers and Humans Apart (CAPTCHA) is nowadays utilized by numerous application service providers in order to defend themselves against automated software agents whose purpose is to degrade the quality of a provided service, whether due to misuse or resource expenditure. CAPTCHA challenges are utilized worldwide by millions of users and thus, it becomes obvious that having usability flaws in the design of CAPTCHA mechanisms, results in unacceptable tradeoffs for the users in terms of time and money. Thus, considering usability aspects in designing CAPTCHA mechanisms becomes nowadays a necessity.

Nowadays, the vast majority of deployed CAPTCHA mechanisms rely on recognition of distorted text with audio as an alternative means for those with vision problems. Research on CAPTCHA design has received significant attention lately. There is a growing demand to enhance both the security and the usability aspects of CAPTCHA mechanisms, aiming to offer high security standards to application service providers and interaction transparency to actual users. Therefore, the design of secure and usable CAPTCHA mechanisms is considered to be a challenging endeavor [1].

There are two main research directions for improving usable security of CAPTCHA mechanisms. The first one is towards breaking and improving existing mechanisms, resulting in the development of design guidelines for increased security [2, 3]. Being focused on security, CAPTCHA mechanisms are currently designed as tasks which are not integrated smoothly in the primary user interaction flow and many attempts exist aiming to provide more usable mechanisms [4-14, 22].

Existing text-based CAPTCHA mechanisms produce challenges with characters solely from the Latin alphabet. This is a “one-size-fits-all” approach that may not provide a viable long-term solution or even a good usability tradeoff, considering that not all Internet users are familiar with this alphabet. This argument is further strengthened bearing in mind that the Internet population consists in its majority of non-native-English speakers¹ and numerous web sites consist of exclusively localized content.

In this context, it is necessary to explore alternative constructs that consider the diversity of Internet users in terms of conventions, cultural, and cognitive backgrounds; tools and resources; and contexts in which CAPTCHA challenges are required to be solved. Towards this direction, we aim to explore the applicability of *localized CAPTCHA* i.e., a system that presents challenges with characters, glyphs, or other symbols in users' native language and cultural context. In contrast, a *Latin-based* one presents challenges using characters only from the Latin alphabet i.e., lowercase and/or uppercase letters from “a” to “z”.

A localized CAPTCHA adapted to users' contextual factors may actually *increase* the usability of a system, in terms of efficiency and effectiveness, while *not decreasing* the security level needed for protecting the system. In this context, the work presented in this paper is primarily driven by the need to apply User Centered Design (UCD) methodologies related to design and evaluation of CAPTCHA mechanisms and contributes towards this direction by investigating whether individual differences, focusing mainly on users' lingual characteristics, affect efficiency and effectiveness on interacting with text-based CAPTCHA mechanisms.

The rest of the paper is structured as follows: first, we review the related works. Subsequently, we describe the context of an empirical study, the methods, and the developed hypotheses. Then, we analyze and discuss the findings of the study. Finally, we summarize and outline the implications of the reported research.

¹<http://www.internetworldstats.com/stats.htm>

2 Related Works

The role and involvement of the real users and their unique characteristics is getting the attention of the research community lately. Based on over a million collected responses, an overall success rate of 96.1% was observed on the ReCAPTCHA service [15]. The service presents garbled words from English scanned books that OCR software failed to recognize. The authors noticed that “*non-English speakers seem to perform slightly worse than English speakers*” based on statistically significant differences in success rates correlated with the origin of the respondent’s IP address [16].

A large-scale study involving over 1,000 volunteers and 21 different text-based CAPTCHA systems revealed among other that “*non-native speakers of English take longer to solve CAPTCHAs and are less accurate on CAPTCHAs that include English words*” [17]. Users needed on average 9.8 seconds to solve a text challenge and between 12 and 25 seconds for an audio challenge. Overall, native-English speakers were substantially faster and slightly more accurate on solving the challenges. We note that the recruitment platform that was used in the study (Amazon Mechanical Turk) is currently available only in the US and thus, it might be the case of bilingual speakers. The study neither evaluated CAPTCHA challenges in languages other than English nor the performance of non-native English speakers.

As part of one accessibility studies, a localized CAPTCHA based on the Urdu language was developed for web-based applications [17]. Urdu is widely spoken in India. A small-scale evaluation of the system involved 50 Urdu-knowing users with no or elementary knowledge of English language. The study found that users were more accurate and solved faster challenges in Urdu rather than in English. Similar results were reported in a recent study with Arab participants [21].

Furthermore, a lab usability study of text-based English CAPTCHA involved 24 participants who spoke Chinese and learned English as a second language [18]. The authors observed that “*When reading a string of alphanumeric characters, the participants spoke the alphabets in English, but the numbers in Chinese.*” This resulted in erroneous responses as for example due to the similar-sounding number “one” in Chinese and letter “e” in English.

Another study investigated users’ perceptions towards localized and Latin-based CAPTCHA challenges [19]. According to the perceptions of the participants, current implementations do not provide an acceptable tradeoff solution with regards to usability. Analysis of responses revealed that even experienced users had severe difficulties in coping with current CAPTCHA challenges: “*every other participant claimed to need two or more tries each time to solve a challenge.*” Furthermore, participants stated to be positively attributed with regards to localized CAPTCHA challenges, especially when visiting web sites with localized content, since their interaction flow tends to be more natural.

A follow-up work examined the relation between the cognitive style of the user and their preference and performance upon interacting with text- and picture-based challenges [20]. The results of the study revealed an existing main effect relating to users’ cognitive styles (Verbalizer or Imager) and a respective performance issue in solving text- vs. picture-based CAPTCHA challenges.

We build on top of these findings aiming to investigate whether individual differences affect preference to localized- vs. Latin-based CAPTCHA challenges. We describe in the next section a study that utilized the developed localized CAPTCHA mechanism, the hypotheses that have been formulated for the purpose of our research together with analysis of interactions and the interpretation of results.

3 Method of Study

3.1 Procedure and research hypotheses

A research prototype of a localized CAPTCHA mechanism was designed and developed aiming to produce challenges of random characters based on different national alphabets. The research prototype was based on an open source CAPTCHA mechanism which has been enhanced as to produce CAPTCHA challenges based on UTF-8 and thus capable to produce challenges in a variety of national alphabets. The research prototype produced challenges in Latin and non-Latin alphabet with a similar level of text distortion. We formulated the following null hypotheses for the purpose of this research:

H_0 . There is no general preference of users towards Latin or non-Latin CAPTCHA challenges, considering also various main effects and interactions with respect to language presentation of the web site and users' national alphabet.

3.2 Experimental design

The aim of the experiment was to examine user preference towards a specific type of CAPTCHA challenge by considering also various main effects and interactions with respect to language presentation of the web site and users' national alphabet. The participants were asked to visit a web page which was placed in front of a web application that announced the students' marks for the course assignments. In this study, the web page content appeared equal randomly in either local language (in Greek) or translated in the English language. The aim was to examine whether the web page content presentation language influenced the user preference towards selecting a specific type of CAPTCHA challenge.

Once a student initiated a session, a random process chose if the student would receive content in English or local language. This choice was fixed for the whole session. The page explained what a CAPTCHA challenge is and presented two example challenges, one in local language and one in English. It then asked the user to select the CAPTCHA type of preference. The placement of the two challenges was equally split: in half cases the localized CAPTCHA example challenge appeared on the left side of the screen and the English one on the right; in the other half this order was reversed. Once the CAPTCHA type choice was made, the participant was faced with the actual CAPTCHA challenge and was asked to provide the correct response. If a correct answer was received, the system displayed the marks for the course assign-

ments or else it challenged again the participant with a new challenge in the same (previously selected) language of preference.

3.3 Demographics of the participants and validity of the study

The participants of the experiment were juniors and seniors of two different academic institutions studying computing science and applications. They were all native non-English speakers. We recruited 384 students in total for this experiment. The age of the students was between 20 and 25 years old with only minor exceptions. In total, 440 different challenges were presented to these students.

The validity of a study is related to its internal and external validity, and ecological factors. There has been an effort to increase ecological validity of the research, since the participants were involved at their own working conditions and physical environments (without the intervention of any experimental equipment or person). Furthermore, participants were asked to solve the CAPTCHA challenge during a real-life task (the primary task of the participants has been to view their course marks). This closely resembles a typical usage scenario related to CAPTCHA usage in which the challenge is as a secondary task. In terms of external validity, we mention that further empirical research with a larger sample is needed aiming to compare the findings of this study with ones performed using other alphabets, more representative samples, and diverse cultural contexts.

4 Study Results and Discussion

In this section, we present and discuss the main findings of the study based on the analysis of the collected responses.

The experimental design explored the actual user preference and the effect of content presentation in this choice. The users visited a web page that contained an example of both a Latin-based and a localized CAPTCHA and asked them which language they prefer to be actually challenged. The page appeared either in local language or in English and the appearance order of the two examples (left-right) was random. For the purpose of the study, we separated the participants in two groups based on their skills with regards to the English language. The control Group (A) consisted of $N=298$ participants who were proficient in English. The experimental group (B) consisted of $N=86$ participants who were not proficient in English.

In Table 1, we summarize the CAPTCHA preferences for each group according to the users' choice. Those participants belonging to Group (A) select to solve the English challenge in the majority of cases (188/298), while those participants belonging to Group (B) select to solve the localized one (72/86). This difference in preference is significant $\chi^2(1, N=384)=58.650$, $p<0.001$.

Furthermore, we examined the effect of the language presentation of the web content towards users' preference for a specific type of CAPTCHA challenge. The choice of the participants in both groups is not affected by the page content language: control group $\chi^2(1, N=298)=0.220$, $p=0.639$ and experimental group $\chi^2(1, N=86)=0.332$, $p=0.564$. The choice is also not affected by the order in which the two example chal-

lenges appear: control group $\chi^2(1, N=286)=1.828$, $p=0.176$ and experimental group $\chi^2(1, N=86)=3.084$, $p=0.079$.

The aforementioned results suggest that there is a main factor between the user preference on CAPTCHA challenges and their familiarity with the Latin alphabet. It appears that, if the users are given a choice, then the users belonging to the experimental Group (B) prefer to solve challenges in their native language, whereas users belonging to the control group (A) prefer to solve the Latin-based CAPTCHA challenges.

Table 1. Group Preference on challenge language by page content.

	Page Content		Total
	Latin	Localized	
Group (A)			
English	87	54	141
Localized	101	56	157
Total	188	110	298
Group (B)			
English	7	42	49
Localized	7	30	37
Total	14	72	86

The aforementioned results are based on objectively measured interaction data and are in line with the results presented in [19], which investigated users' preference in solving localized vs. Latin-based CAPTCHA challenges through a questionnaire based survey. Bearing in mind that these recent studies followed different methodological approaches (objective vs. subjective self-reporting data) it seems promising to consider user's diversity with regards to their lingual characteristics in the design of CAPTCHA challenges, as nowadays CAPTCHA designs and implementations clearly fail to consider in their iteration design cycles such individual characteristics of the users as they produce challenges with characters solely from the Latin alphabet.

Simultaneously, such an approach would have as well some positive side effects, especially in use cases in which the users interact with a web application in native language content presentation, as it is preferable to have a uniform presentation, including the actual site content and the CAPTCHA challenge. Thus, we argue that enhancing existing Latin-based CAPTCHA mechanisms with the capability to provide localized text-based challenges would result to a better user experience.

5 Conclusions

The purpose of this paper is to presented results of an empirical study with the aim to gain knowledge for enhancing usability and user experience related to CAPTCHA

challenges as an effort to apply User Centered Design (UCD) approaches related to CAPTCHA mechanisms. We argue that for designing more usable mechanisms, research in this area should partially move its focus away from the technical security issues towards understanding the users and developing approaches which can be applied for increasing usability and offering a better user experience to users who are engaged in solving CAPTCHA challenges.

In this context, given that current CAPTCHA designs provide challenges solely in Latin alphabet and considering the diversity of World Wide Web users with regards to their lingual characteristics we developed a CAPTCHA mechanism capable to produce challenges of different alphabets. We utilized the developed CAPTCHA mechanism in the frame of an ecological valid experimental design aiming to examine users' preference towards Latin and localized CAPTCHA challenges. The analysis of results was mainly focused on users' preference towards localized and Latin-based CAPTCHA challenges. However, we also looked at the language presentation factor of the web page in which the CAPTCHA mechanism was embedded aiming to examine if this factor has a main effect towards users' preference to a specific type of CAPTCHA challenge. We recruited overall 384 non-native English-speaking users aiming to investigate the aforementioned factors and revealed that there is a statistical significant preference of users towards using specific type of CAPTCHA challenges depending on their profession in using the Latin-alphabet.

A future research direction, aiming to examine in wider perspective human factors on CAPTCHA designs, is to investigate preference and performance factors in localized and Latin- based CAPTCHA challenges in comparative cross-national studies. Bearing in mind that CAPTCHA tasks are not embedded smoothly into the primary task of a user (e.g. a user not familiar with the Latin alphabet performs a registration on a local content web site that requires to solve a Latin – based CAPTCHA challenge) and that every second millions of users world wide are solving CAPTCHA challenges, studies like the reported one could have an impact on making these tasks more affordable as even a small improvements will have massive benefits for users.

6 References

1. Yan, J. and A.S. El Ahmad. "Usability of CAPTCHAs or usability issues in CAPTCHA design". In Proceedings of ACM SOUPS 2008, pp. 44–52.
2. Burzstein, E., M. Martin, and J.C. Mitchell. "Text-based CAPTCHA Strengths and Weaknesses". In Proceedings of ACM CSS 2011, pp. 125-138.
3. El Ahmad A.S., J. Yan, and M. Tayara. "The Robustness of Google CAPTCHAs". Newcastle University Tech. Report CS-TR-1278, September 2011.
4. Matthews P. and C.C. Zou. Scene tagging: image-based CAPTCHA using image composition and object relationships. In Proc. of the 5th ACM Symposium on Information, Computer and Communications Security (CCS 2005), ACM, New York, NY, USA. pp. 345-350.
5. He, P., Y. Sun W. Zheng, and X. Wen. "Filtering short message spam of group sending using CAPTCHA". In Proceedings of the First International Workshop on Knowledge Discovery and Data Mining (WKDD 2008), pp. 558-561.

6. Kluever K.A. and R. Zanibbi. "Balancing usability and security in a video CAPTCHA". In Proceedings of ACM SOUPS 2005, ACM, New York, NY, USA, pp. 14:1-14:11.
7. Zhang, H. , X. Wen, P. He, and W. Zheng. "Dealing with telephone fraud using CAPTCHA". In Proceedings of the 2009 Eight IEEE/ACIS International Conference on Computer and Information Science, IEEE Computer Society, Washington, DC, USA. pp. 1096-1099.
8. Datta, R., Li, J., Wang, J.Z., 2005. IMAGINATION: A robust image-based CAPTCHA generation system, in: Proceedings of the 13th annual ACM international conference on Multimedia, ACM, New York, NY, USA. pp. 331-334.
9. Cui, J., Zhang, W., Peng, Y., Liang, Y., Xiao, B., Mei, J., Zhang, D., Wang, X., 2010. A 3-layer dynamic CAPTCHA implementation, in: Education Technology and Computer Science (ETCS), 2010 Second International Workshop on, pp. 23-26.
10. Lin, R., Huang, S.Y., Bell, G.B., 2011. A new CAPTCHA interface design for mobile devices, in: Lutteroth, C., Shen, H. (Eds.), Australasian User Interface Conference (AUIC 2011), ACS. pp. 3-8.
11. Banday, M., Shah, N., 2009. Image Fip CAPTCHA. ISeCure, The ISC International Journal of Information Security, 105-123.
12. Ho, C.J., Wu, C.C., Chen, K.T., Lei, C.L., 2011. DevilTyper: a game for CAPTCHA usability evaluation. *Comput. Entertain.* 9, 3:1-3:14.
13. Holman, J., Lazar, J., Feng, J.H., D'Arcy, J., 2007. Developing usable CAPTCHAs for blind users, in: Proceedings of the 9th international ACM SIGACCESS conference on Computers and accessibility, ACM, New York, NY, USA. pp. 245:246.
14. Ross, S.A., Halderman, J.A., Finkelstein, A., 2010. Sketcha: a CAPTCHA based on line drawings of 3D models, in: Proceedings of the 19th international conference on World wide web, ACM, New York, NY, USA. pp. 821-830.
15. von Ahn, L., B. Maurer, C. McMillen, D. Abraham, and M. Blum. reCAPTCHA: Human-Based Character Recognition via Web Security Measures. *Science*, 321(5895):1465, 2008.
16. Bursztein, E., S. Bethard, C. Fabry, J.C. Mitchell, and D. Jurafsky. "How good are humans at solving CAPTCHAs? A large scale evaluation". In Proceedings of the 2010 IEEE Symposium on Security and Privacy (SP '10), IEEE Computer Society, Washington, DC, USA. pp. 399-413.
17. Banday, M.T. and N.A. Shah. "Challenges of CAPTCHA in the accessibility of Indian regional websites". In Proceedings of the Fourth Annual ACM Bangalore Conference (COMPUTE '11), ACM, New York, NY, USA. pp. 31:1-31:4.
18. Lee, Y.L., Hsu, C.H. "Usability study of text-based CAPTCHAs". *Displays* 32, pp. 81-86.
19. Fidas, C., A.G. Voyiatzis, and N.M. Avouris. "On the necessity of user-friendly CAPTCHA". In Proceedings of the 2011 Annual Conference on Human Factors in Computing Systems (CHI'11), pp. 2623-2626.
20. Belk, M., C.A. Fidas, P. Germanakos, and G. Samaras. "Do Cognitive Styles of User Affect Preference and Performance related to CAPTCHA?" In Proceedings of the 2012 Annual Conference on Human Factors in Computing Systems (CHI 2012), pp. 1487-1492.
21. Ashraf Khalil, Salam Abdallah, Soha Ahmed, Hassan Hajjdiab. "Script Familiarity and Its Effect on CAPTCHA Usability: An Experiment with Arab Participants", *International Journal of Web Portals*, 4(2), 74-87, April-June 2012.
22. Te-En Wei, Jeng, A.B.; Hahn-Ming Lee. "GeoCAPTCHA - A novel personalized CAPTCHA using geographic concept to defend against 3rd Party Human Attack", *Performance Computing and Communications Conference (IPCCC)*, 2012 IEEE 31st International, pp. 392- 399.