SUBLIMINAL WARNINGS – A NEW APPROACH TO CHANGE USERS' BEHAVIOR

Authors: Mini Zeng, Feng Zhu, Sandra Carpenter

Presenter: Mini Zeng
mzeng@ju.edu
Computer Science Department
Jacksonville University
Jacksonville, USA

DR. MINI ZENG

- Assistant Professor of Computing Science and Cybersecurity, Jacksonville University
- Research Interest:
  - Human Computer Interaction, Eyetracking and EEG, Machine learning and data mining, Pervasive Security and Privacy, Software Security, Intrusion Detection
- Publication and Activities:
  - Systems and methods of providing gaze-based notifications, U.S. Patent App. Patent No. US 10,162, 651, B1,
  - Subliminal Warnings: Utilizing the High Bandwidth of Nonconscious Visual Perception. International Conference on Persuasive Technology 2022
  - Secure Coding in Five Steps." Journal of Cybersecurity Education, Research and Practice 2021
  - Designing warnings to reduce identity disclosure. International Journal of Human–Computer Interaction, 2018,

INTRODUCTION

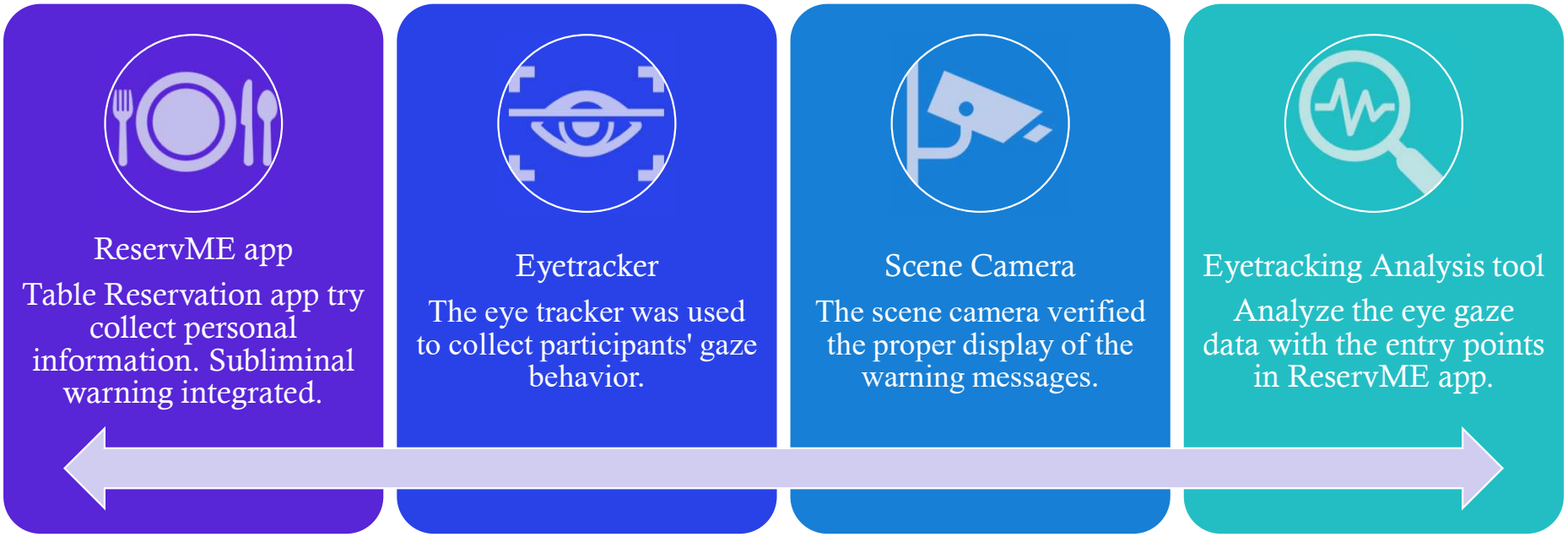
Background:

- There are 17 million identity theft victims every year in the U.S. Unfortunately. While people are very concerned about their privacy, they do not protect their personal information well and unnecessarily expose their information online.
- Researchers believe one of the fundamental problems is the lack of attention paid to warnings. So, recent warning designs try to grab users' attention by discontinuing their primary tasks. Other researchers claimed that too much exposure to warnings and interruptions would make users quickly habituate to warnings and feel "warning fatigue".

What we do:

- We utilize the non-conscious processing ability for warnings instead of competing for conscious attention.
- We bypass the other stages of the C-HIP model and directly use the stimulus-response mapping to trigger the safe behavior.
- We want to evaluate the impact of cybersecurity warnings via non-conscious processing, specifically using subliminal messages.

EXPERIMENT ENVIRONMENT



METHOD

TABLE 1 PILOT STUDY CONDITIONS

Condition	Subliminal Prime	Duration for each display	Number of Displays	Display
1	Text message “Fake it”	50 ms	Five Times	At the top of Address textbox
2	Text Message “Fake it”	50 ms	Once	At the top of Address textbox
3	Icon: Yellow Triangle	50 ms	Five Times	At the right of “Next” button
4	Icon: Yellow Triangle	50 ms	Five Times	At the right of Address textbox

TABLE 2 PILOT STUDY RESULT: PERCENTAGE OF PARTICIPANTS DISCLOSING PERSONAL INFORMATION

Subliminal Prime	Condition	First name	Last name	Email	Phone Number	Address	Zip code
Message	1	71%	57%	43%	29%	43%	57%
	2	80%	80%	20%	20%	25%	20%
Icon	3	100%	100%	80%	80%	80%	80%
	4	100%	100%	100%	80%	100%	100%

EXPERIMENT PROCEDURE

- We used an eye-gaze based verification system with an eye tracker and a scene camera and a post-experiment questionnaire for evaluation.
- We recruited 58 participants on campus for the experiment.
- The design of this experiment has two conditions: the control condition (with no warning) and the subliminal warning condition (with the subliminal message to "fake it", display 50ms on top of the address line one).
- We target the street address for not disclosure based on an assumption that the street address would be more sensitive than another field (zip code, state, etc) for a table reservation app.
- The participant signed the consent form and disclaimer first. The disclaimer stated that the restaurant reservation app was developed by a third-party software company and the purpose of this study is to evaluate the useability of the software.

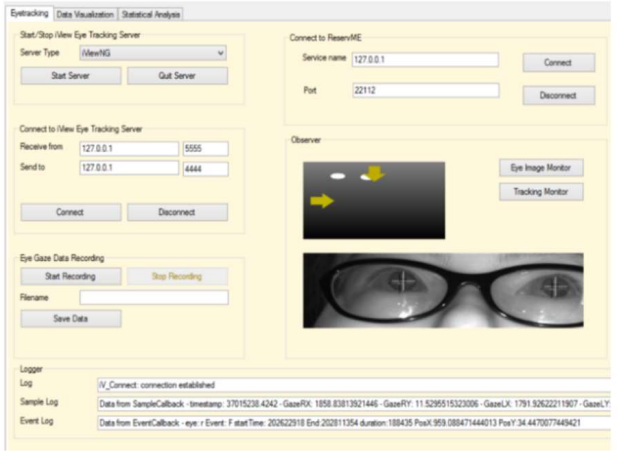


Figure 1. Integrated software environment for the experiment

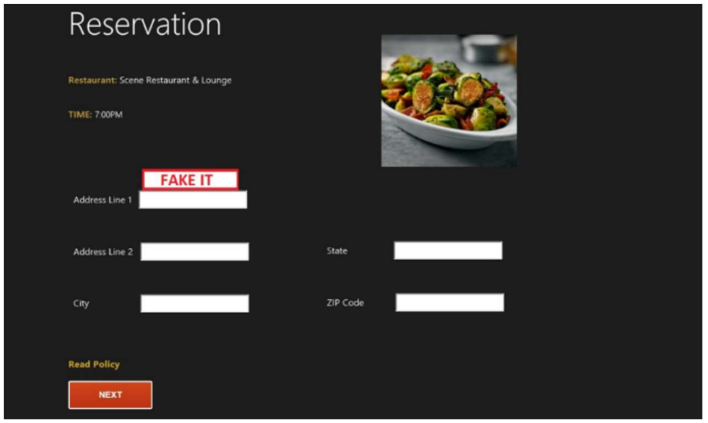


Figure 2. A subliminal warning was shown when a user mouse-clicked the input textbox.

EXPERIMENT RESULT

- Figure 3 shows a participant's eye gaze during the display of the subliminal warning message. The colored dots represent eye-gaze locations with 4ms intervals that overlay on top of each (the lighter the color, the more recent the eye gaze location).
- Table 3 shows subliminal warning experiment results for the two experimental conditions. Probability values (p-values) less than 5% (bold) indicate statistically significant differences between the control and subliminal message conditions.

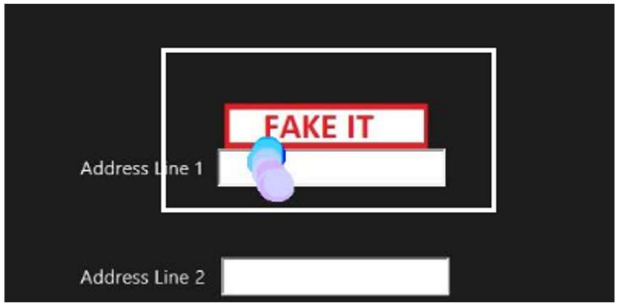


Figure 3. A participant's eye gaze information when a subliminal warning shows up

TABLE 3. SUBLIMINAL WARNING EXPERIMENT RESULTS

	Did not input information				**Faked information**				**Exposed real information**			
	Con-trol	**Sub-liminal**	**P-value**	**Odds ratio**	**Con-trol**	**Sub-liminal**	**P-value**	**Odds ratio**	**Con-trol**	**Sub-liminal**	**P-value**	**Odds ratio**
Address	1	5	**0.035**	6.30	5 (29)	10 (23)	**0.048**	2.78	24	13	**0.004**	0.22 (4.62)
City	0	2	0.068	∞	3 (30)	7 (26)	**0.065**	3.00	27	19	**0.017**	0.23 (4.26)
State	0	2	0.068	∞	3 (30)	5 (26)	0.192	1.95	27	21	0.065	0.33 (3.00)
Zip code	0	2	0.068	∞	3 (30)	9 (26)	**0.019**	4.26	27	17	**0.004**	0.17 (5.82)
Email	0	4	**0.016**	∞	2 (30)	10 (24)	**0.003**	7.78	28	14	**<0.001**	0.07 (12.1)
Phone	2	5	0.095	3.04	3 (28)	9 (23)	**0.019**	4.26	25	14	**0.003**	0.20 (5.00)

EXPERIMENT RESULT

TABLE 4. PARTICIPANTS' EXPLANATIONS OF WHY THEY DID OR DID NOT PROVIDE THEIR PERSONAL INFORMATION.

Reasons for Providing Truthful Information		
	Control	**Subliminal**
Information is not sensitive or it is not risky to disclose.	7	11
Trust in the app or experimental context.	12	6
Information is needed by the app or experiment.	4	2
Information is already known.	2	2
Habit.	4	0
Reasons for Falsifying Personal Information		
	Control	**Subliminal**
Information is sensitive or risky to disclose.	4	11
Distrust the app/experimental context.	1	1
Information is not needed by the app/experiment.	1	1
Told to "fake it."	0	3

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.581[a]	.338	.298	1.33455

ANOVA

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	45.393	3	15.131	8.496	.000[b]
	Residual	89.051	50	1.781		
	Total	134.444	53			

Coefficients

Model		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	t	Sig.
1	(Constant)	2.406	1.295		1.858	.069
	DistanceToSubliminalMsg	1.543	.386	.475	4.003	.000
	Attitudes	-.091	.065	-.175	-1.403	.167
	Concerns	-.042	.039	-.132	-1.086	.283

Figure 4. Liner regression results. The relationships between behavior, privacy attitudes, concerns, and eye gaze locations.

SUMMARY

The display of the subliminal warnings

- The average display duration of the warning message was 84ms, although we set the display duration for 50ms (see Section 4.2 for how we measured the durations). For twenty-seven participants, the warning message lasted 56ms (5 frames) to 101ms (9 frames).

Were Participants Consciously Aware of the Warning Message?

- Six out of twenty-eight participants in the subliminal warning condition could report the subliminal warning message in one way or another. Three of them recalled the exact action suggested, "FAKE IT;" two remembered the message "falsify" which was semantically correct, and the other one ("red boarder") seemed not to have processed the warning message at the semantic level. Unlike the other five, this participant provided accurate information for all identity elements.

Identity Disclosure Behavior

- More participants in the subliminal warning condition did not input their information for all six identity elements than those in the control condition. For participants who input their information, those in the subliminal condition were much more likely to fake their identity elements than those in the control condition.

What the Eye Tracking Data Show?

- In twenty out of twenty-four cases, participants' eye gazes were near the subliminal warning message as shown in the white rectangle area (up to 50 pixels away in X or Y coordinates from the "FAKE IT" message) as shown in Figure 3.

How Participants Explained their Disclosure Behavior

- For example, "I feel uncomfortable sharing this information" or "this information is safe to give." The second most frequent reason for disclosure was related to trust/distrust of the app or the experimental context (n = 20 responses), such as "I trust apps like this." Note that only half as many participants in the warning condition, compared to the control condition, indicated that they trusted the app/ experiment.

CONCLUSION AND FUTURE WORK

Conclusion

- Design a subliminal warning by suggesting safe behavior using stimulus–response mapping model.
- When an app or a website collects a large amount of identity information, we can suggest users to “fake it.” Cognitively more straightforward to trigger the safe behavior .
- We conducted a pilot study and tested two categories of warnings (message and icon) to guide the design of the subliminal warning.
- We used eye tracking and scene camera recording to verify the display duration of the subliminal warning and users' attention during the experiment.
- The result of the experiment showed that the subliminal warning with the suggested response could effectively reduced disclosure of identity information.

Future Work

- Implement the application solution of the subliminal warning. We envision that an application implementation of subliminal warning could be developed as a third-party application or as a web browser plug-in.
- We plan to extend the experimental study with multiple times displays of subliminal warnings, different warning words, duration, colors, and background of the message.
- We will also apply other statistical analyses such as the omnibus test to test different parameters and Bonferroni correction to mitigate family-wise errors.
- Compare the effectiveness of the different strategies and to find their limitations.

REFERENCES

[1] J. Sobey, R. Biddle, P. C. van Oorschot, and A. S. Patrick, “Exploring user reactions to new browser cues for extended validation certificates,” in *European Symposium on Research in Computer Security*, 2008, pp. 411–427.

[2] A. Wiesmann, A. Stock, M. Curphey, and R. Stirbei, “A guide to building secure web applications and web services,” *The Open Web Application Security Project*, 2005. http://www.i-pi.com/Training/BSS/OWASPGuide2.0.1.pdf (accessed May 19, 2022).

[3] K. Krol, M. Moroz, and M. A. Sasse, “Don’t Work. Can’t Work? Why It’s Time to Rethink Security Warnings,” *Risk and security of internet and systems (CRiSIS), 2012 7th International conference*, pp. 1–8, 2012.

[4] A. Vance, D. Eargle, J. L. Jenkins, C. B. Kirwan, and B. B. Anderson, “The fog of warnings: how non-essential notifications blur with security warnings,” in *Fifteenth Symposium on Usable Privacy and Security (SOUPS 2019)*, 2019, pp. 407–420.

[5] S. Carpenter, F. Zhu, M. Zeng, and M. Shreeves, “Expert sources in warnings may reduce the extent of identity disclosure in cyber contexts,” *International Journal of Human--Computer Interaction*, vol. 33, no. 3, pp. 215–228, 2017.

[6] M. S. Wogalter, “Communication-Human Information Processing(C-HIP)Model,” in *Handbook of Warnings*, M. Wogalter, Ed. Mahwah,NJ, USA, 2006, pp. 51–62.

[7] B. Anderson, T. Vance, B. Kirwan, D. Eargle, and S. Howard, “Users aren’t (necessarily) lazy: using NeuroIS to explain habituation to security warnings,” *Thirty Fifth International Conference on Information Systems*, pp. 1–15, 2014.

[8] D. Akhawe and A. P. Felt, “Alice in warningland: a large-scale field study of browser security warning effectiveness,” *Proceedings of the 22nd USENIX Security Symposium*, pp. 257–272, 2013.

[9] C. Bravo-Lillo, L. F. Cranor, J. Downs, and S. Komanduri, “Bridging the gap in computer security warnings: A mental model approach,” *IEEE Security and Privacy*, vol. 9, no. 2, pp. 18–26, 2011, doi: 10.1109/MSP.2010.198.

[10] S. Dehaene, J. P. Changeux, L. Naccache, J. Sackur, and C. Sergent, “Conscious, preconscious, and subliminal processing: a testable taxonomy,” *Trends in Cognitive Sciences*, vol. 10, no. 5, pp. 204–211, 2006, doi: 10.1016/j.tics.2006.03.007.