SeRIM 2025

Venice, Italy | March 4th 2025

Replication Study: Cross-Country Evaluation of the Recognition-Based Graphical Authentication Scheme in AR and VR Environments

Naheem Noah, Peter Mayer, Sanchari Das

BACKGROUND

AR/VR HMDs present unique authentication challenges, especially shoulder-surfing *(Zhang et al., 2023).*

Traditional authentication methods (e.g., passwords, PINs) are often ill-suited for AR/VR *(Lebeck et al., 2018).*

Graphical password schemes offer improved usability, memorability, and shoulder-surfing *(Düzgün et al., 2022).*

The Things scheme is a recognition-based graphical authentication system *(Düzgün et al., 2022).*

Original evaluation focused on AR only (HoloLens 2) with 16 German participants.

RESEARCH OBJECTIVES

01

Replicate the Things graphical authentication study in a new cultural and technological context

02

Expand evaluation from AR-only (HoloLens) to include VR (Valve Index)

03

Gather quantitative insights on effectiveness, efficiency, usability, security, cognitive load

04

Gather qualitative feedback on user experiences, comfort, and preferences

Things Scheme

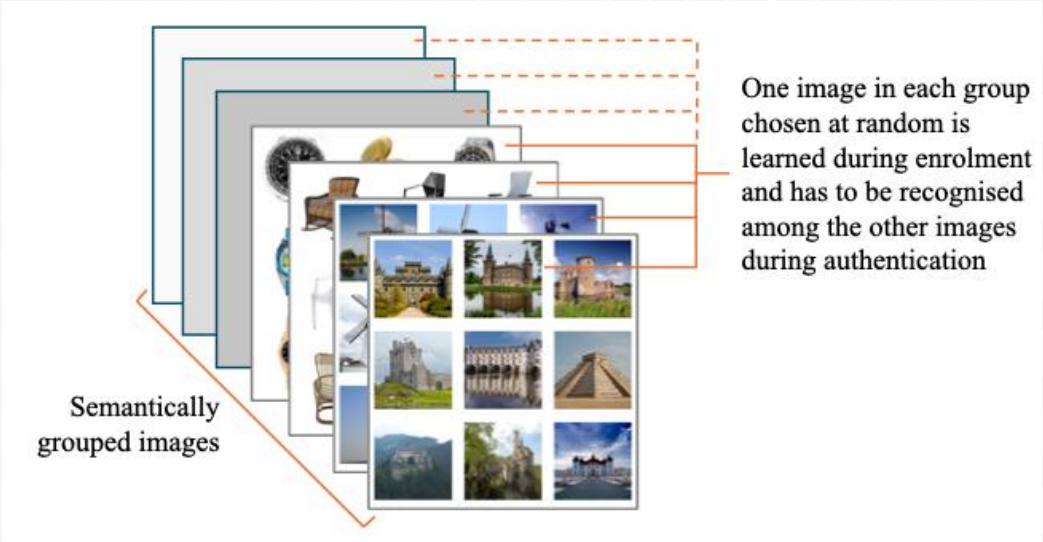

Figure 1. The Things authentication scheme: Images are semantically grouped and displayed sequentially in a grid. One image per group is part of the user's password. See Figure 3 for our AR/VR implementation.

METHODOLOGY

Things Configuration

Semantic Grouping of Images

Password Structure

Password Space Design

Random Assignment

METHODOLOGY

Study Design

Between-Subjects Design

Prescreening, in-lab orientation, demonstration, enrollment, authentication attempts, and post-study survey

5 system-assigned images

Session lasted ~30 minutes, in-person

IRB approved

METHODOLOGY

Quantitative Measures

- Effectiveness
- Efficiency
- Perceived Security
- Perceived Usability
- Perceived Task Load

Qualitative Feedback

- Overall Experience
- Specific Challenges Encountered
- Level of Comfort
- Security & User Friendliness
- Additional Thoughts or Suggestions

METHODOLOGY

Figure 2. Overview of Participants Recruitment & Research Methodology.

METHODOLOGY

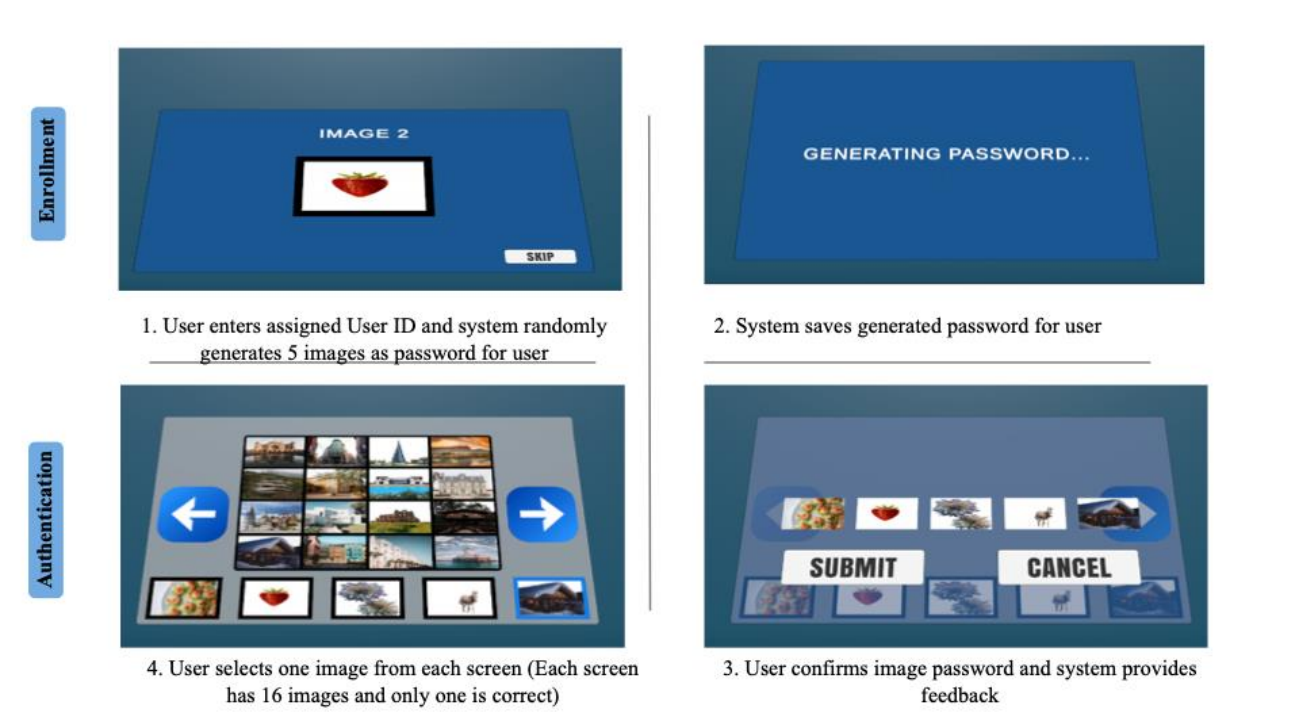

Figure 3. User VR Environment Showing the Enrollment and Authentication Process in the Things Scheme

FINDINGS

Table 1. Participant demographics and characteristics

Demographics Info	*Düzgün et al.*	HoloLens	Valve Index	Both
Participants	16	16	16	32
Age Range (years)				
18-24	62.5%	37.5%	25%	31.25%
25-30	31.25%	43.75%	56.25%	50%
31-40	0%	6.25%	12.5%	9.38%
41-50	6.25%	12.5%	6.25%	9.38%
Gender				
Male	56.25%	75%	87.5%	81.25%
Female	43.75%	25%	12.5%	18.75%

FINDINGS

Table 1. Participant demographics and characteristics

Educational Background	*Düzgün et al.*	HoloLens	Valve Index	Both
Masters	No info	37.5%	43.75%	40.63%
Bachelors	No info	43.75%	43.75%	43.75%
Doctoral	No info	6.25%	0%	3.13%
High School	No info	6.25%	12.5%	9.38%
Diploma	No info	6.25%	0%	3.13%
Prior Usage of AR/VR Headsets	75%	50%	68.75%	59.38%
Future Use of AR/VR Headset	91.67%	81.25%	93.75%	87.5%
Ownership of AR/VR Headset	6.25 %	12.5%	6.25%	9.38%
Prior password entry in AR/VR Headset	0%	18.75%	12.5%	15.63%

FINDINGS

Table 2. Details of scheme evaluations based on effectiveness, security, and usability. * = Statistically significant

	Düzgün et al.	Our Study Overall	HoloLens	Valve Index
Participants	16	32	16	16
Effectiveness & Efficiency				
SUS Score	74	72.81	70.47	75.16
Average Enrollment Duration	62.21s (SD = 24.76)	25.23s (SD = 1.54)	25.38s (SD = 1.63)	25.01s (SD = 1.63)
5-second Interval Rating	3.81	4.03	3.81	4.25
Average Authentication Duration	32.2s (SD = 9.39)	36.45s (SD = 36.80)	49.85s (SD = 47.53)*	23.06s (SD = 9.56)*
Average Successful Authentication Duration	-	25.23s (SD = 1.54)	29.87s (SD = 6.42)	19.02s (SD=4.50)
Overall Success Rate	90%	79% (SD = 0.39)	73% (SD = 0.46)	85% (SD = 0.33)
Participants Success in three iterations	75%	75%	68.75%	81.25%

FINDINGS

Table 2. Details of scheme evaluations based on effectiveness, security, and usability. * = Statistically significant

	Düzgün et al.	Our Study Overall	HoloLens	Valve Index
Perceived Security				
Overall System Security	3.19 (SD = 1.01)	3.94 (SD = 1.18)	3.90 (SD = 1.22)	3.94 (SD = 1.14)
System Security against Shoulder-surfing	3.94 (SD = 1.18)	4.4 (SD = 0.74)	4.25 (SD = 0.75)	4.4 (SD = 0.74)
Perceived Usability				
Easy to Use	4.00 (SD = 1.12)	3.81 (SD = 1.18)	3.5 (SD = 1.32)	4.13 (SD = 0.93)
Easy to Remember	4.31 (SD = 0.77)	3.88 (SD = 1.39)	3.81 (SD = 1.55)	3.94 (SD = 1.20)
Fast Login Process	3.19 (SD = 1.38)	3.94 (SD = 1.32)	3.69 (SD = 1.26)	4.19 (SD = 1.33)
Willingness to Use in the Future	3.25 (SD = 1.09)	4.06 (SD = 1.06)	3.88 (SD = 1.17)	4.25 (SD = 0.90)

FINDINGS

Table 2. Details of scheme evaluations based on effectiveness, security, and usability. * = Statistically significant

	Düzgün et al.	Our Study Overall	HoloLens	Valve Index
Perceived Task Load				
Mental Demand	-	2.63 (SD = 1.02)	2.56 (SD = 1.12)	2.69 (SD = 0.92)
Physical Demand	-	2.28 (SD = 1.15)	2.56 (SD = 1.22)	2 (SD = 1)
Temporal Demand	-	2.47 (SD = 0.83)	2.5 (SD = 0.87)	2.44 (SD = 0.79)
Successful Accomplishment	-	4.41 (SD = 1.17)	4.19 (SD = 1.29)	4.6 (SD = 0.99)
Level of Hard work	-	2.5 (SD = 0.97)	2.56 (SD = 1.12)	2.44 (SD = 0.79)
Insecurity or Stress-level	-	1.66 (SD = 0.92)	1.94 (SD = 0.97)	1.38 (SD = 0.78)

FINDINGS

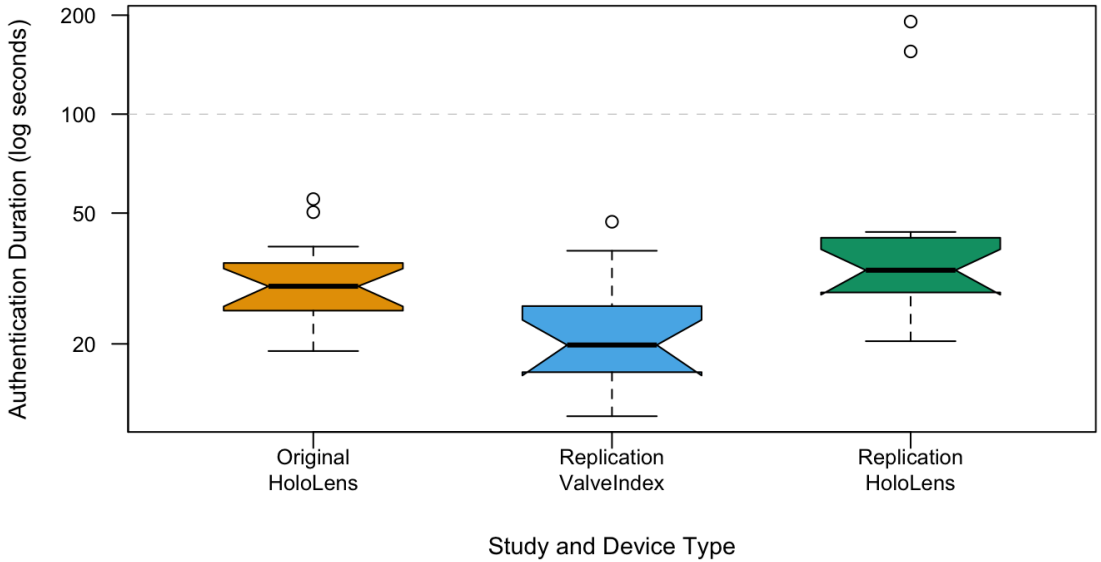

Figure 3. User VR Environment Showing the Enrollment and Authentication Process in the Things Scheme

FINDINGS: QUALITATIVE RESPONSE

	VALVE INDEX	HOLOLENS
OVERALL EXPERIENCE	*"I found the scheme easy to use overall. Once I got going with the system, it was very quick to enter my password, much quicker than It would have been for me to type. out a pin or password on a virtual keyboard."*	*"Clicking each item and focusing it was the hardest part. It was having a difficult time sensing clicks and focusing."*

FINDINGS: QUALITATIVE RESPONSE

	VALVE INDEX	HOLOLENS
SPECIFIC CHALLENGES ENCOUNTERED	*“Nothing specific. The only thing that might make it better would be a slightly longer display time for the images during registration, but that’s all I can think of.”*	*“The main challenge was obviously remembering the images. Since I am used to traditional, [[textbased]] password schemes, I tried to map the images to certain words and assign initials to each image. By assigning those initials I was able to [[remember]] the password for the duration of the experiment. It is highly likely that I will forget it after a while.”*

FINDINGS: QUALITATIVE RESPONSE

	VALVE INDEX	HOLOLENS
SECURITY & USER FRIENDLINESS	*"It seems like it would be just as secure as the other options. I found it easier to use than a password or pin, but more difficult to use than biometrics."*	*"In terms of user-friendliness, and visual appeal, it was definitely very pretty. In comparison it to other authentication methods, I find it to be weaker than them. Maybe because other schemes have certain factors that allow it to be computationally unfeasible or require having a secondary physical device for two-factor auth, and I don't see any such thing being introduced here."*

DISCUSSION: Comparing Replicated Study

SUS scores were comparable: 72.81 (replication) vs. 74 (original)

Average enrollment time dropped from 62.21s (original) to 25.23s (replication)

Replication had longer average authentication time (36.45s) vs. original (32.2s), but successful attempts were faster

Success rate decreased: 79% vs. 90%, though 75% of participants succeeded in all attempts in both studies

Participants in the U.S. rated overall system security higher (3.94) vs. original study's 3.19, especially for shoulder-surfing resistance

Willingness to reuse the system was higher in the replication (4.06) compared to the original (3.25)

DISCUSSION: Comparing HoloLens and Valve Index Study

Valve Index outperformed HoloLens in SUS score and Authentication success rate

Faster and More Efficient Interaction in VR

Higher Perceived Usability in Valve Index

Controller-based interaction on Valve Index was seen as more intuitive and less fatiguing

Comparable Password Memorability

Security Perception Remained High on Both

DISCUSSION: Cognitive and Physical Demands

Overall Workload Was Low to Moderate

Valve Index Was Less Demanding

High Perceived Task Success in Valve Index

Low Levels of Frustration

HoloLens users experienced more strain, whereas Valve users reported a more comfortable interaction

DISCUSSION: Cultural and Regional Differences

Higher Perceived Security in U.S. Participants

Greater Willingness to Reuse in U.S.

Cultural Attitudes May Influence Trust

University Sample Limits Cultural Conclusions

IMPLICATIONS

VR Offers Superior Usability

Interaction Design Matters

Strong Potential for Shoulder-Surfing Resistance

Cognitive and Physical Demand Are Manageable

Cultural Context Affects Security Perception

Customizability May Boost Memorability

LIMITATIONS

Used HoloLens 1 instead of the HoloLens 2 used in the original study

Participant pool was 81.25% male

Only two platforms (HoloLens and Valve Index) were tested

Study duration was limited to a single session

FUTURE WORK

Explore the Things scheme on emerging devices

Investigate user-selected images for password

Conduct long-term evaluations to assess memorability, fatigue, and behavior over time

Expand participant diversity to understand how different user groups interact with AR/VR authentication schemes

THANK YOU!

scan for pre-print

naheem.noah@du.edu | mayer@imada.sdu.dk | sdas35@gmu.edu

@naheemnoah