

Supplementary File: Can I Borrow Your ATM? Using Virtual Reality for (Simulated) In Situ Authentication Research

Florian Mathis*
University of Glasgow
University of Edinburgh

Kami Vaniea†
University of Edinburgh

Mohamed Khamis‡
University of Glasgow

A STRUCTURED INTERVIEW: AUTHENTICATION SCENARIO

Note that questions marked with a * are asked on a 5-point Likert scale (*strongly disagree to strongly agree*).

1. Please walk us, in detail, through the task you have just experienced.
2. What was your main goal? Please explain why.
3. What were the tasks that were required to achieve that goal?
4. What (if any) is the difference between withdrawing cash at a real-world bank ATM and what you have just experienced?
5. *If participants' cash withdrawal was not successful (e.g., wrong PIN):* What were the main difficulties when trying to withdraw the amount of cash we asked you to withdraw?
6. "While completing the task, I felt I was part of a laboratory study." *
7. "I was aware of the experimenter during the task." *
8. "The experimenter's presence impacted my performance negatively." *
9. "The experimenter's presence impacted my behaviour." *
10. "I found that recalling the PIN made it more challenging to complete the other cash withdrawal steps." *
11. "I found that the other cash withdrawal steps made it more challenging to recall the correct PIN." *

B SEMI-STRUCTURED INTERVIEW (END OF STUDY)

We used a semi-structured interview approach at the end of the study. The following questions were used to roughly ask the same questions to all participants but due to the nature of a semi-structured interview approach the questions differed across the participants.

1. Could you please walk us through your ranking on: "Which experience did you perceive as most similar to using an ATM in the real world?"
2. How did you feel about interacting with the ATM in the real world? What was easy and/or challenging?
3. How did you feel about interacting with the ATM in virtual reality? What was easy and/or challenging?

*e-mail: florian.mathis@glasgow.ac.uk

†e-mail: kvaniea@inf.ed.ac.uk

‡e-mail: mohamed.khamis@glasgow.ac.uk

4. Please consider the experienced environment and a real-world environment where you are standing in front of an ATM. What would be different to what you have just experienced in:

(a) our real-world part of the study?

(b) our VR part of the study?

5. Did the amount of cash you had to withdraw impact your authentication behaviour? If so, how?
6. Do you regularly shield your PIN entry when using an ATM in the real world?
7. *If yes to 6):* How do you shield your PIN entry?
8. Did you shield your PIN entry in the study? Why? Why not?
9. *If yes to 8):* How did you shield your PIN entry in the study?
10. What do you think this study is about?

C RAW NASA-TLX SCORES FOR EACH SUBDIMENSION

Table 1 shows the raw NASA-TLX scores for each subdimension.

D STATISTICAL ANALYSIS: F-RATIOS FOR AUTHENTICATION TIME, NUMBER OF CORRECTIONS, AND NUMBER OF ERRORS

Table 2 shows the F-ratios for participants' authentication times, number of corrections, and number of errors.

Table 1: The table shows the dimensions of the NASA-TLX scores. We did not perform post-hoc tests on the level of each dimension due to the lack of significance of the overall mean raw NASA-TLX scores. Scores represent the mean and the stdev.

NASA-TLX	(1) RW Lab	(2) RW ATM	(3) VR Lab	(4) VR ATM	(5) VR ATM Public		
Mental (ColorPIN only)	44.00 (29.18)	59.50 (28.85)	41.75 (28.38)	58.25 (26.38)	61.50(25.70)	No significant main effects on the overall NASA-TLX scores.	
Mental (ATM + ColorPIN)	n/a	39.00 (28.31)	n/a	50.00 (32.71)	58.25 (27.81)		
Physical (ColorPIN only)	9.25 (9.91)	10.25 (11.34)	19.00 (20.10)	13.25 (19.38)	19.00 (18.68)		
Physical (ATM + ColorPIN)	n/a	15.75 (18.05)	n/a	22.25 (21.12)	23.50 (18.38)		
Temporal (ColorPIN only)	33.25 (31.32)	24.50 (24.89)	25.25 (26.90)	25.00 (27.88)	30.75 (27.31)		
Temporal (ATM + ColorPIN)	n/a	25.50 (24.59)	n/a	22.75 (26.05)	39.50 (22.13)		
Performance (ColorPIN only)	34.75 (39.48)	31.25 (35.53)	27.25 (33.30)	23.00 (32.65)	33.00 (33.44)		
Performance (ATM + ColorPIN)	n/a	37.25 (35.62)	n/a	31.50 (35.11)	33.25 (34.14)		
Effort (ColorPIN only)	42.25 (27.36)	44.00 (32.58)	43.00 (30.47)	44.25 (26.80)	55.25 (24.47)		
Effort (ATM + ColorPIN)	n/a	44.75 (25.57)	n/a	50.00 (25.45)	51.25 (24.02)		
Frustration (ColorPIN only)	27.25 (21.24)	36.50 (28.86)	34.00 (28.09)	34.5 (27.88)	40.75 (25.85)		
Frustration (ATM + ColorPIN)	n/a	37.00 (26.29)	n/a	34.50 (25.59)	39.50 (22.63)		
Overall Workload Score (ColorPIN only)	31.79 (30.48)	34.33 (32.03)	31.71 (29.49)	33.04 (30.91)	40.04 (30.03)		p > 0.05
Overall Workload Score (ATM + ColorPIN)	n/a	33.21 (28.60)	n/a	35.17 (30.29)	40.88 (27.78)		p > 0.05

Table 2: The table shows the statistical analysis, including means, stdevs, F-ratios, effect size, and p-values of participants' authentication times (in seconds), number of corrections, and number of errors. $p < 0.05$ highlighted.

Measure (Two-way RM ANOVA)	(1) RW Lab	(2) RW ATM	(3) VR Lab	(4) VR ATM	Context (Lab/ATM)	Environment (RW/VR)	Context×Environment	p<0.05
Authentication Time	13.28 (7.76)	16.57 (14.01)	20.89 (8.33)	23.85 (25.32)	$F(1,49) = 0.149, p = 0.70, \eta_p^2 = 0.003$	$F(1,49) = 27.00, p < 0.05, \eta_p^2 = 0.36$	$F(1,49) = 0.313, p = 0.58, \eta_p^2 = 0.006$	1-3,2-4
Number of Corrections	0 (0)	0.45 (1.07)	0.40 (0.73)	0.30 (0.90)	$F(1,57) = 0.418, p = 0.52, \eta_p^2 = 0.007$	$F(1,57) = 0.269, p = 0.61, \eta_p^2 = 0.005$	$F(1,57) = 0.516, p = 0.48, \eta_p^2 = 0.009$	n/a
Number of Errors	0.60 (1.11)	0.65 (1.07)	0.55 (0.92)	0.40 (0.92)	$F(1,57) = 0.420, p = 0.52, \eta_p^2 = 0.007$	$F(1,57) = 0.157, p = 0.69, \eta_p^2 = 0.003$	$F(1,57) = 0.650, p = 0.42, \eta_p^2 = 0.01$	n/a
Measure (One-way RM ANOVA)	(3) VR Lab	(4) VR ATM	(5) VR ATM Public	Context	p<0.05			
Authentication Time	20.89 (8.33)	23.85 (25.32)	25.55 (13.73)	$F(2,33) = 3.676, p < 0.05, \eta_p^2 = 0.18$	n/a			
Number of Corrections	0.40 (0.73)	0.30 (0.90)	0.20 (0.68)	$F(2,38) = 0.73, p = 0.49, \eta_p^2 = 0.04$	n/a			
Number of Errors	0.55 (0.92)	0.40 (0.92)	0.75 (0.99)	$F(2,38) = 1.40, p = 0.259, \eta_p^2 = 0.07$	n/a			