Better the Devil You Know: A User Study of Two CAPTCHAs and a Possible Replacement Technology

Kat Krol, <u>Simon Parkin</u>, M. Angela Sasse

University College London (UCL), UK

Motivation

- CAPTCHAs are a hurdle to completion of tasks
- Collectively a massive draw on time
- Company approached us claiming a zero-effort biometric
- Assessed viability alongside other solutions
  - Efficiency, effectiveness, satisfaction (ISO9241-11)

Technologies

- **reCAPTCHA** – squiggly characters!

- **PlayThru** – a themed drag-and-drop game

- **NoBot** – user positions their face in view of the device camera. Pictures are taken and compared to a database
  - An authentication solution, used here for human verification

Study setup

- Mock ticket buying website – a primary task
- Participants asked to purchase tickets three times
  - Browse → select → verify → enter provided details
- Verifying they are human was part of check-out
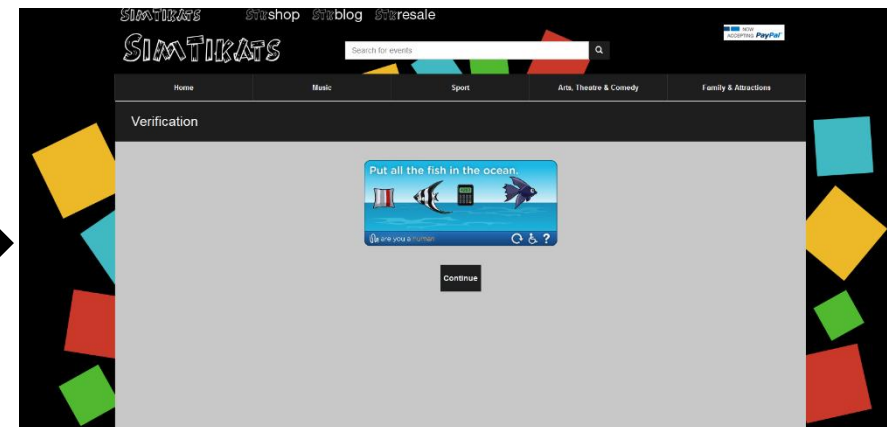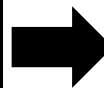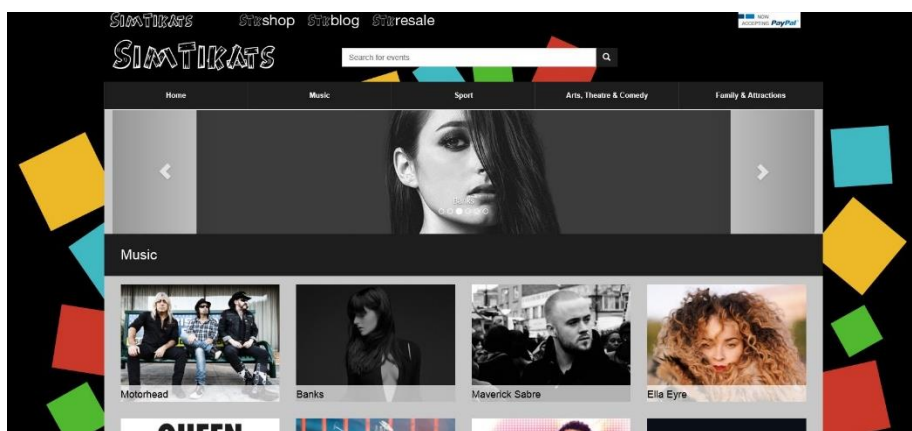
Study conditions

1. Mixed – complete ticket purchases on a laptop
   – Order randomised for each participant
2. NoBot Laptop – three successive purchases verified with NoBot
3. NoBot Tablet

- Mixed condition allowed comparison
- Repeated use assesses learnability
- Followed by interview, NASA-TLX, adjectives, contexts

Measurements

- For each purchase verification
  - Time (measure time on dedicated verification page)
- For each technology
  - Workload: NASA-TLX
  - Descriptive adjectives
  - Context where use considered appropriate
  - Ranking
- Semi-structured interviews
  - Transcripts coded

Participants

- 11 participants in the pilot
- 87 valid participants in the main study (57f, 30m)
  - 27 used NoBot on a laptop (3x)
  - 31 used NoBot on a tablet (3x)
  - 29 verified using reCAPTCHA, PlayThru and NoBot on a laptop (1x)

Time for NoBot conditions
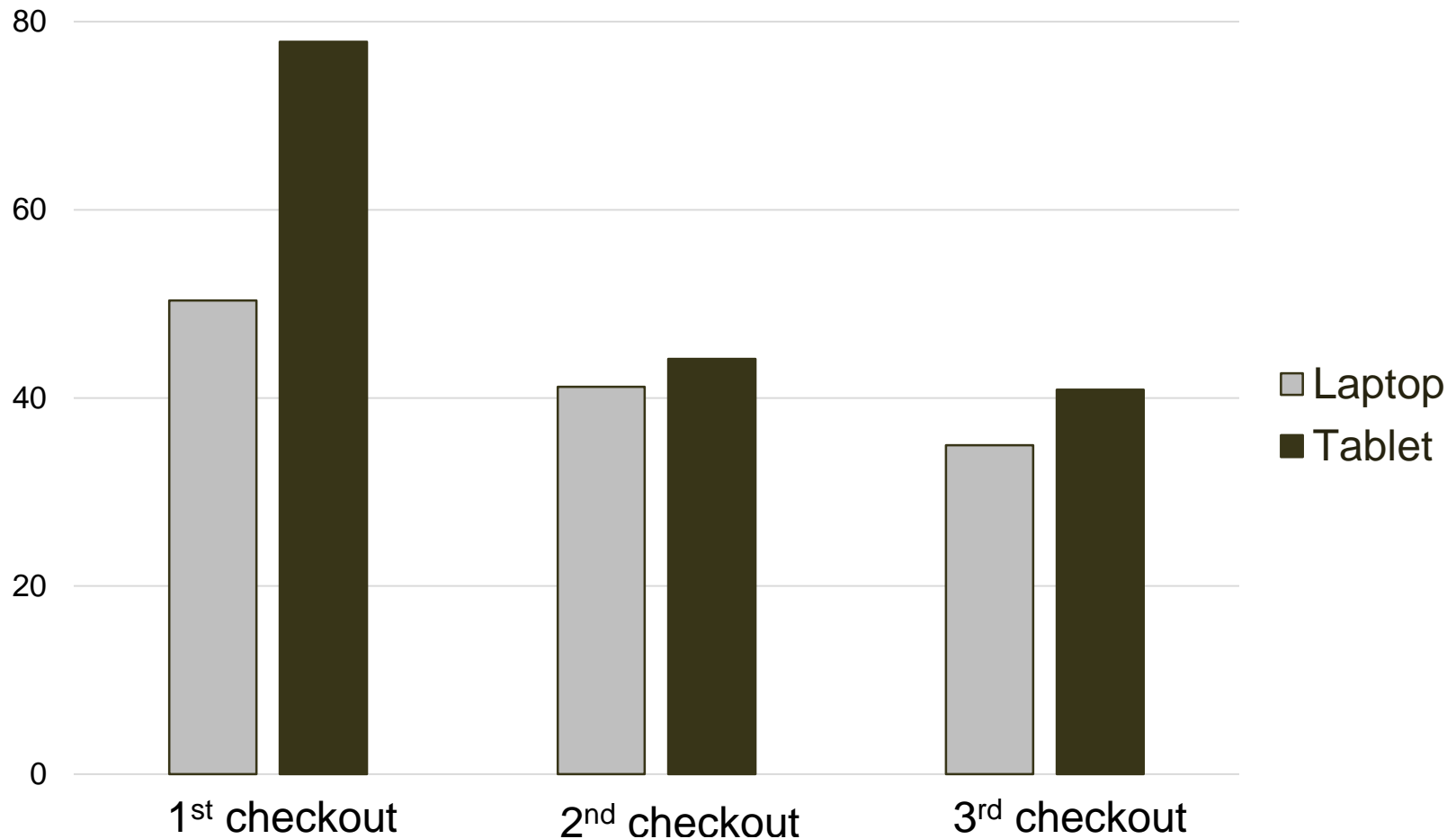
Average time in seconds taken to verify using NoBot across devices and trials

Time in the comparison condition
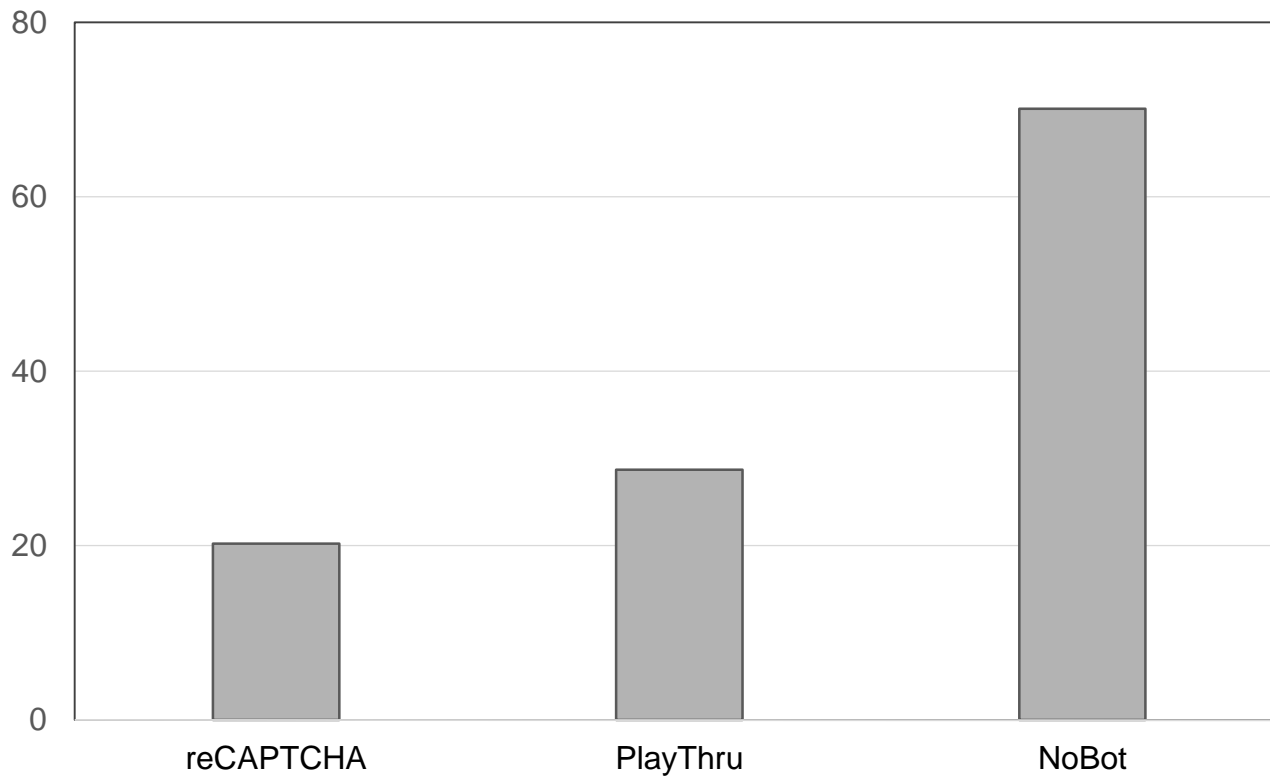
Average time in seconds taken to verify using the three mechanisms

Workload (NASA TLX): laptop vs. tablet

TLX aspect	Laptop	Tablet
Mental demand	4.2	4.8
Physical demand	**2.9**	**7.2**
Temporal demand	5.6	5.1
Performance	7.2	7.9
Effort	5.4	7.0
Frustration	5.4	5.4
Overall workload	29.9	37.4

Workload (NASA TLX): three mechanisms

TLX aspect	reCAPTCHA	PlayThru	NoBot
Mental demand	10.75	9.24	8.89
Physical demand	4.33	8.45	8.34
Temporal demand	4.62	3.08	6.47
Performance	3.62	5.46	5.51
Effort	8.12	8.21	6.46
Frustration	**8.73**	**6.31**	**11.7**
Overall workload	48.4	37	46.7

Adjectives chosen to describe NoBot

Laptop		Tablet	
Effortless	13	Effortless	11
Fast	9	Intuitive	10
Intuitive	8	Slow	9
Weird	7	Easy to use	8
Easy to use	6	Acceptable	8

Top five adjectives chosen by participants to describe their experience with NoBot.

Adjectives chosen to describe all mechanisms

reCAPTCHA		PlayThru		NoBot	
Normal	14	Acceptable	14	Unpredictable	9
Acceptable	13	Exciting	9	Weird	9
Effortful	9	Effortless	7	Creepy	8
Easy to use	8	Intuitive	7	Fast	8
Predictable	7	Great	8	Effortful	6

Top five adjectives chosen by participants to describe their experience with reCAPTCHA, PlayThru and NoBot.

Contexts

Context	reCAPTCHA	PlayThru	NoBot
Contributing to an online forum	59	15	24
Buying tickets online	93	79	55
Browsing for plane tickets	76	55	45
Checking in for flights online	86	54	52
Topping up Oyster card	66	69	31
Bidding on items in eBay	66	69	31
Logging in to Facebook from a different computer	66	66	34

Mixed condition - ranking of mechanisms
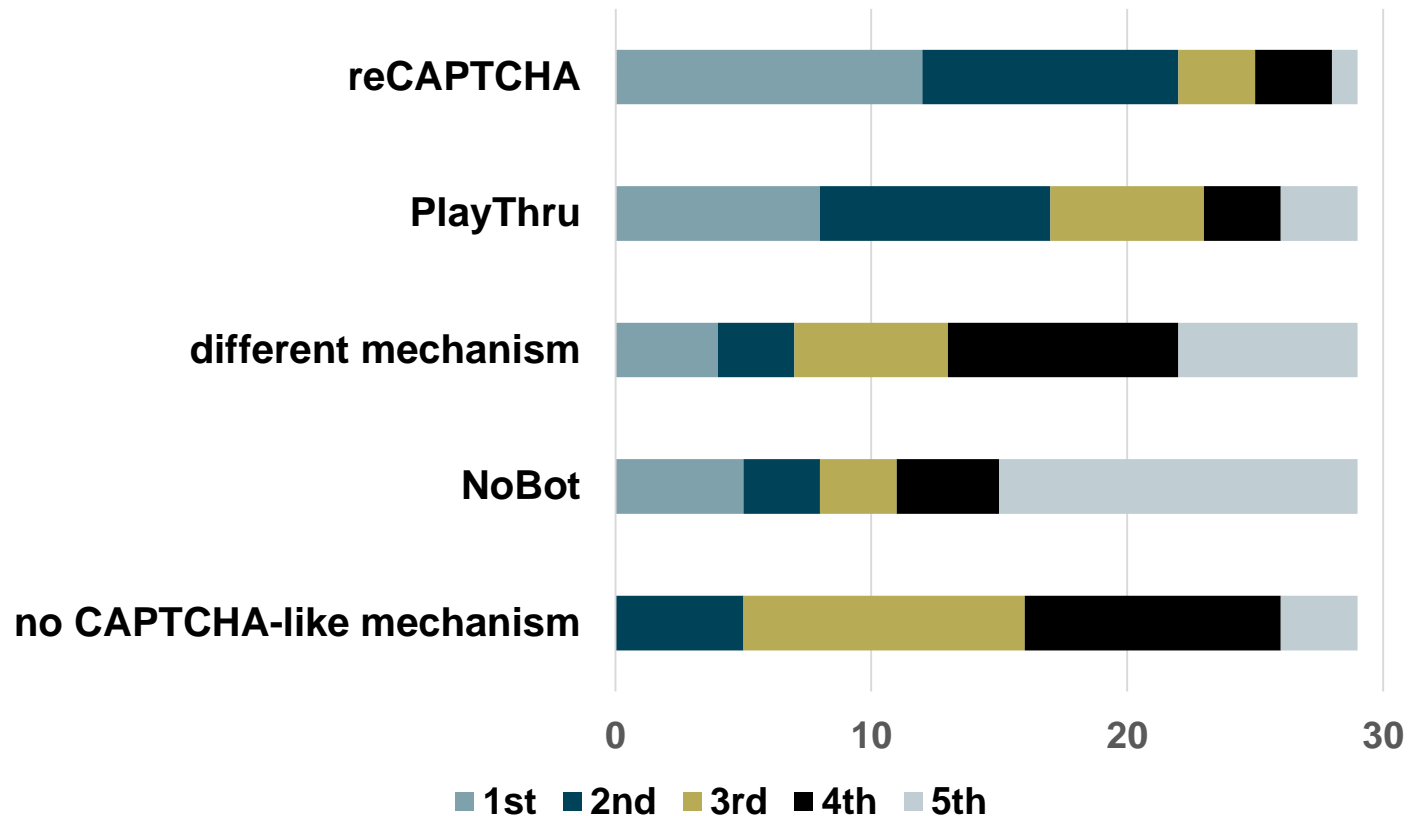
Views on text-based CAPTCHAs

- Most frequently used words: 'annoying' (44), 'frustrating' (20) and 'hate' (10)

- *"I actually hate all of these because sometimes I really can't see properly, I'm not a robot but I just can't see, you know, and I keep refreshing the CAPTCHA. It's quite troublesome."* (PL08)

Views on PlayThru

- *"It disturbed my usual routine, it required some kind of effort which actually I didn't exhort that much, I just guessed through it and it was correct what I did. If not, I imagine it would have been very frustrating."* (PM03)

Views on NoBot: Privacy

- Many participants thought that NoBot's purpose was to identify them

- Participants asked who'd see their images and what inferences they would make: *"I liked it less because I don't like putting my face on the Internet; taking a picture of me and who knows what I'm buying."* (PM05)

Views on NoBot: Security

- 32 participants described scenarios where the security of their images would be compromised like a database hack

- *"I would be concerned about anyone hacking into this system and then I have all these personal details including my face."* (PL16)

Views on NoBot: Reliability

- *"like checking in for flights, I would not use NoBot because there might be some error and I could not check in for my flight. Then I would use the traditional because it's easier, I can refresh the images and try again, so I think it would be faster. So if there is something urgent, I would not use it [NoBot]."* (PL11)

NoBot – less effort since "you just stay there"

"[For PlayThru] I still need to read the sentence to know what they want me to do. I need to think and for NoBot, I just stay here and it's done." (PM01)

Conclusions

- reCAPTCHA – a necessary evil
- PlayThru – entertainment factor, unsuitable for serious activities
- NoBot – engaging, but use of images a thorny issue
  - Seen in a better light when assessed in isolation
- Long-term goal to build an assessment framework