The usability canary in the security coal mine: A cognitive framework for evaluation and design of usable authentication solutions

Brian Glass\textsuperscript{1} Graeme Jenkinson\textsuperscript{2} @gcjenkinson Yuqi Liu\textsuperscript{1} M. Angela Sasse\textsuperscript{1} Frank Stajano\textsuperscript{2}

\textsuperscript{1}University College London (United Kingdom) \\
\textsuperscript{2}University of Cambridge (United Kingdom)

EuroUSEC 2016-07-18
A usability canary

*It's as dark as a dungeon way down in the mine*
—Merle Travis
The Quest to Replace Passwords

Bonneau et al. 2012

- Evaluated two decades of proposed password replacement schemes
- 25 usability, deployment and security benefits
- No one scheme better in all dimensions
Authentication doesn’t exist in a vacuum

The impact of a given security measure is a function of:

- When it occurs in the user’s workflow
- What functions of the brain it loads
- What the user was meant to be doing before and after
Modelling business process

- A representation of the set of steps
- Tasks that can be performed at each step
- Hard constraints that enforce partial ordering of tasks
- Soft constraints that capture the costs of switching tasks
Airline self-service checkin kiosk

Background

Airline self-service checkin kiosk

Model

Model validation

Conclusions

Airline self-service checkin kiosk

LANG

AIRL

BKRF

FRBN

LIQH

DIMH

STSO

STSR

EXBG

CFRM

PRLT

PRBP

AUTH
When a person switches from one task to another task, the brain must reorganize and reallocate cognitive resources to ensure an efficient transition.
Operationalizing the check-in task

| From                          | Visual working memory (VWM) | Procedural memory (PM) | Declarative recall (DR) | Semantic recognition (SR) | Episodic recognition (ER) |
|-------------------------------|-----------------------------|------------------------|-------------------------|---------------------------|---------------------------|
| To                            | VWM | PWM | DR  | SR  | ER  |
| VWM                           | 0   | 0.495 | 0.495 | 0.495 | 0.157 |
| PWM                           | 0.495 | 0   | 0.495 | 0.699 | 0.699 |
| DR                            | 0.495 | 0.495 | 0   | 0.482 | 0.482 |
| SR                            | 0.495 | 0.842 | 1.078 | 0   | 0.433 |
| ER                            | 0.307 | 0.842 | 1.078 | 0.354 | 0   |

Table: Costs of switching between tasks utilizing different cognitive mechanisms, given as Cohen’s $d$ effect sizes.
## Operationalizing the check-in task

| Code  | Primary cognitive resource          | Modality       | Voluntary? | Familiarity | Complexity |
|-------|------------------------------------|----------------|------------|-------------|------------|
| LANG  | Semantic recognition               | Touchscreen    | No         | 5           | 1          |
| AIRL  | Episodic recognition               | Touchscreen    | No         | 5           | 1          |
| BKRF  | Visual working memory              | Touchscreen QW-ERTY | No         | 3           | 3          |

**Table:** Properties of the check-in kiosk tasks. Familiarity and complexity are on a scale from 1 (low) to 5 (high).
Constraint satisfaction problems
## Cognitive Psychology

### Constraint satisfaction problems

| Task                                      | Language 1 | Language 2 | Language 3 | Language 4 |
|-------------------------------------------|------------|------------|------------|------------|
| Select language                           | Select language | Select language | Select language | Select language |
| Select airline                            | Select airline | Select airline | Select airline | Select airline |
| Check liquids                             | Check liquids | Check liquids | Check liquids | Check liquids |
| Booking reference                         | Booking reference | Booking reference | Booking reference | Booking reference |
| Check forbidden items                      | Insert payment card | Passport info | Password | Password |
| Select return seat                        | Buy extra bag | Select return seat | Check forbidden items | Check forbidden items |
| Check luggage size                        | Select return seat | Check luggage size | Select outbound seat | Select outbound seat |
| **Passport scan**                         | Check luggage size | Check forbidden items | Check luggage size | Check luggage size |
| Buy extra bag                             | Check forbidden items | Buy extra bag | Buy extra bag | Buy extra bag |
| Confirm                                   | Confirm | Confirm | Confirm | Confirm |
| Print boarding pass                       | Print boarding pass | Print boarding pass | Print boarding pass | Print boarding pass |
| Select outbound seat                      | Select outbound seat | Select outbound seat | Select return seat | Select return seat |
| Print luggage tag                         | Print luggage tag | Print luggage tag | Print luggage tag | Print luggage tag |

| Cost | 5.53 | 5.88 | 8.18 | 8.42 |

**Table:** Optimal task ordering of the self-service check-in using different authentication mechanisms.
Simulated kiosk

- Optimal group $N = 40$
  $(\text{Age}_{\text{Mean}} = 26.6, \text{Age}_{SD} = 7.2, 28$ females)

- Pessimal group $N = 20$
  $(\text{Age}_{\text{Mean}} = 29.1, \text{Age}_{SD} = 13.5, 15$ females)

- On completion participants filled out a subjective satisfaction questionnaire
Completion times

- Significant main effects of Condition ($F_{1,55} = 4.82, p = 0.03$) and Experience ($F_{1,55} = 5.01, p = 0.03$)
- Significant main effect of Repetition ($F_{2,110} = 81.0, p < 0.001$)
Completion times by experience

Significant interaction of Repetition and Experience ($F_{2,110} = 5.09, p = 0.01$)
User satisfaction

Directionally in favor of the Optimal ordering, the satisfaction ratings were not statistically significantly.
Self service checkin kiosk results

Ordering task

Airline Checkin Procedure

Which language?
Which airline?
Forbidden materials?
END

Drag all tasks below into the desired order below...

Airline Checkin tasks

Booking reference
Print luggage tag?
Outbound Seat
Buy extra bag?
Confirm

Reprint boarding pass?
Hand luggage size?
Liquids >100ml?

Submit
Self service checkin kiosk results

Ordering distance

Expert group (n=17), software/web settings (n=33), or both (n=6), with 5.2 mean years of experience (SD = 6.0)
Conclusions

- The cost of a security measure is a function of its relationship to the user’s actual goal.

- The impact of task switching has been operationalized using results from experimental psychology.

- Participants performed better and were more satisfied with the model’s optimally order interface.

- The model’s optimal ordering was more similar to suggested orderings of professional designers.
| Questions? |
|------------|
|           |