RepliSEC

Is FIDO2 the Kingslayer of User Authentication? A Comparative Usability Study of FIDO2 Passwordless Authentication

Agenda

- Topic of research
- Original paper
- Related research
- Study design
- Data analysis

Study

RepliSEC

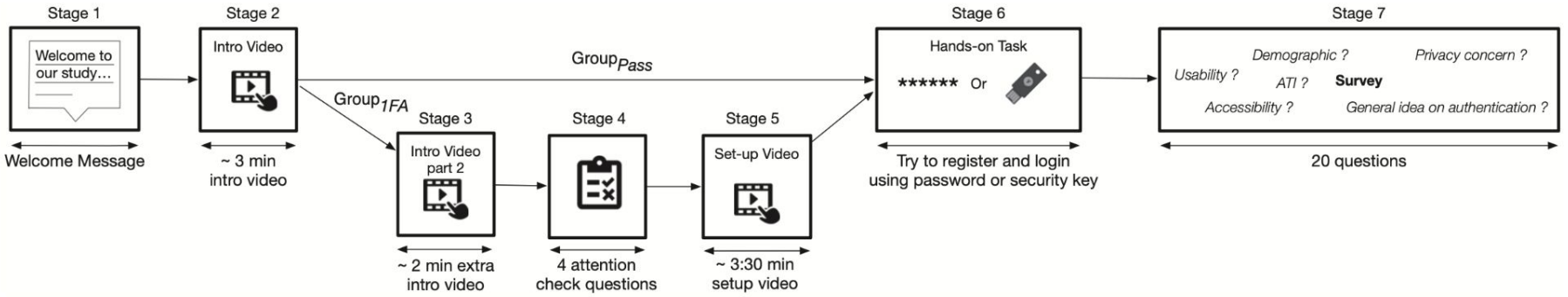
- We want to replicate an existing research
- Research in question
  "Is FIDO2 the Kingslayer of User Authentication? A Comparative Usability Study of FIDO2 Passwordless Authentication"
- How usable is FIDO2 (Passkeys) as a passwordless authentication option?

Related research

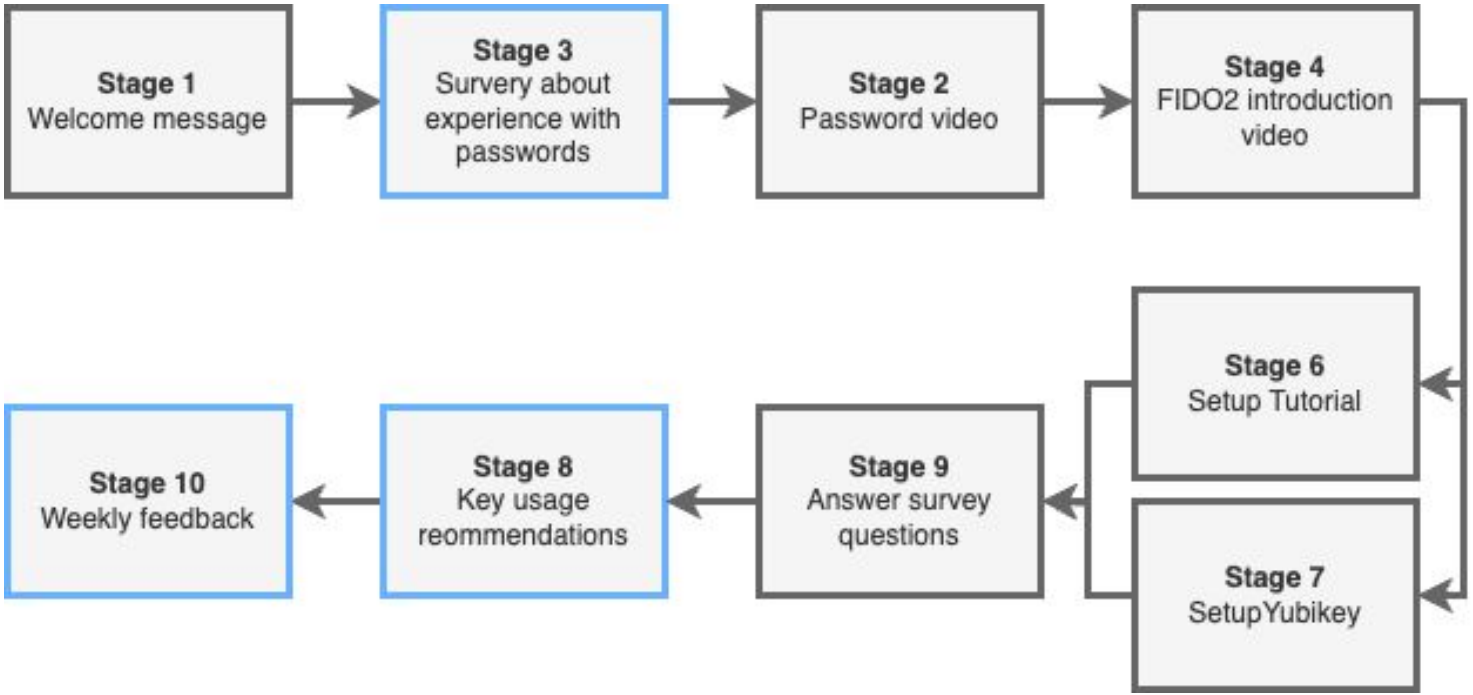
- Florian M. Farke, Lennart Lorenz, Theodor Schnitzler, Philipp Markert, & Markus Durmuth (2020). “You still use the password after all” – Exploring FIDO2 Security Keys in a Small Company. In Sixteenth Symposium on Usable Privacy and Security (SOUPS 2020) (pp. 19–35). USENIX Association.
- Stephane Ciolino, Simon Parkin, & Paul Dunphy (2019). Of Two Minds about Two-Factor: Understanding Everyday FIDO U2F Usability through Device Comparison and Experience Sampling. In Fifteenth Symposium on Usable Privacy and Security (SOUPS 2019) (pp. 339–356). USENIX Association.
- Joshua Reynolds, Trevor Smith, Ken Reese, Luke Dickinson, Scott Ruoti, & Kent E. Seamons (2018). A Tale of Two Studies: The Best and Worst of YubiKey Usability. 2018 IEEE Symposium on Security and Privacy (SP), 872-888.

Explanation of existing paper

The study design can be summarized in this flow diagram



Our Study Design



In Detail

- Participants were colleagues, friends and family who had Gmail/Amazon account using their own laptops and could understand english
- They watched the introduction videos from the original paper
- They filled out the password survey containing SUS, Acceptance, and ATI
- They watched a Passkey setup video and followed the steps in parallel
- They filled-out a survey about the experience with the setup also containing SUS, Acceptance, and ATI
- They were given advices on Yubikey usage and encouraged to try somewhere else
- They filled out a weekly survey in the following three weeks
- On completion they received the used Yubikey as compensation

Major differences to the study

- Using real websites instead of fake websites (Google / Amazon)
- Using the current implementation of FIDO2 passwordless authentication which are Passkeys
- Passkeys require extra PIN
- Within-Group design
- 3-Week-long study to test real-life use-case

Results

Collected data

- Demographic questions
- Password and key setup surveys
  - Acceptance (9 questions)
  - System Usability Scale (10 questions)
  - Affinity for technology interaction (9 questions)
  - Privacy Concern (4 questions)
  - Technical Problems (2 questions)
  - One open-Ended question
- Weekly (first and second week)
  - How many times did they use the Yubikey
  - Did they use it with other services
  - Did they have problems with it
- Final (third week)
  - Same questions as weekly
  - Acceptance + SUS

Problems

- Study took 6 weeks, so we didn't have the time to analyse all data
- Study was running during christmas so most participants didn't have the opportunity to thoroughly use the Yubikey
- One of the ATI-Scale questions was cut short in the survey and we had to point that out to the participants
- Sadly Limesurvey didn't store the tokens to the submissions for the final survey (it did for all others)
- Small sample size: 10 people (in paper 94)
- Small variation in participants

Hypothesis

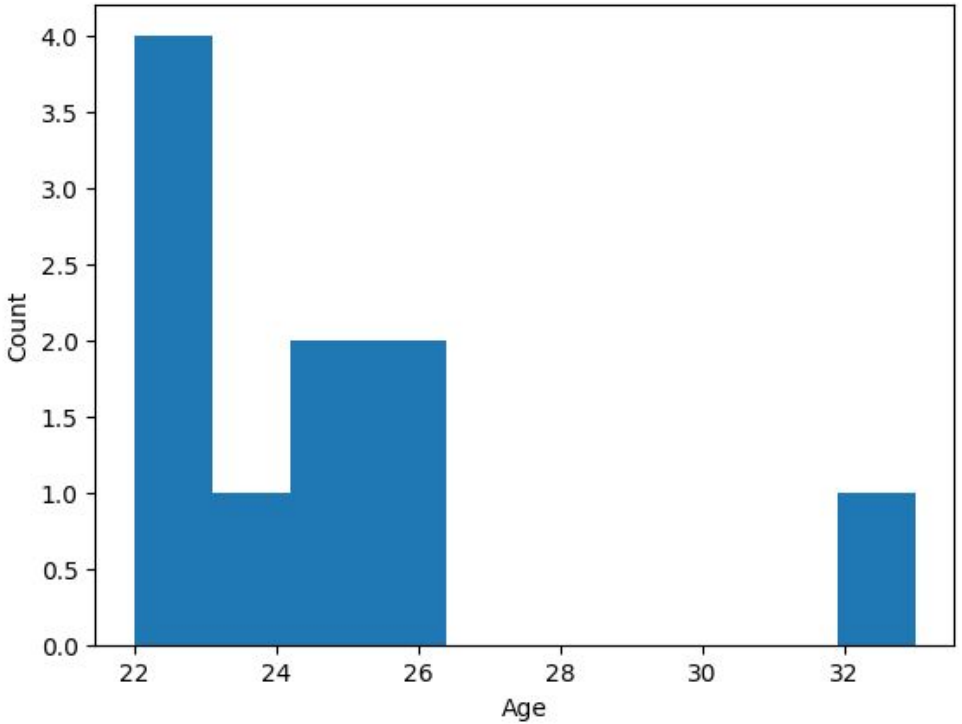
H1: FIDO2 passwordless authentication has a higher usability than traditional password-based authentication

H2: FIDO2 passwordless authentication and the traditional password-based method differ in their acceptance.

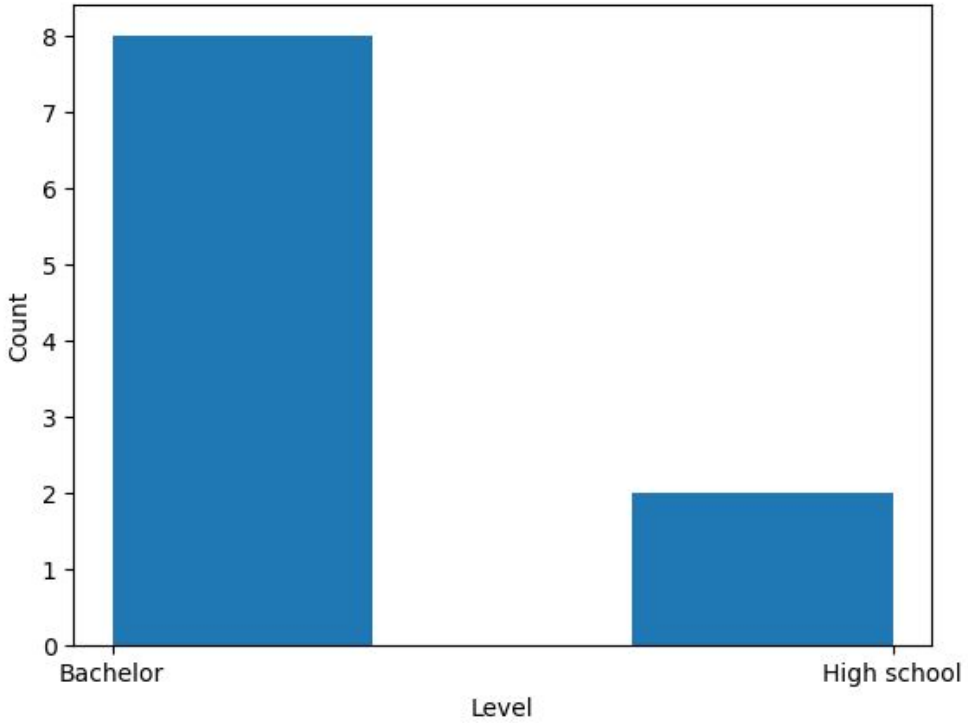
Analysis

- Acceptance and SUS scales' questions were converted to their respective scores using the standard methods
- Values from Two Surveys were compared using paired T-Tests
- ATI scale questions were converted to their respective scores
- Demographic data was plotted as histograms
- Weekly surveys data was also plotted
- Still-pending: Open-Ended questions analysis

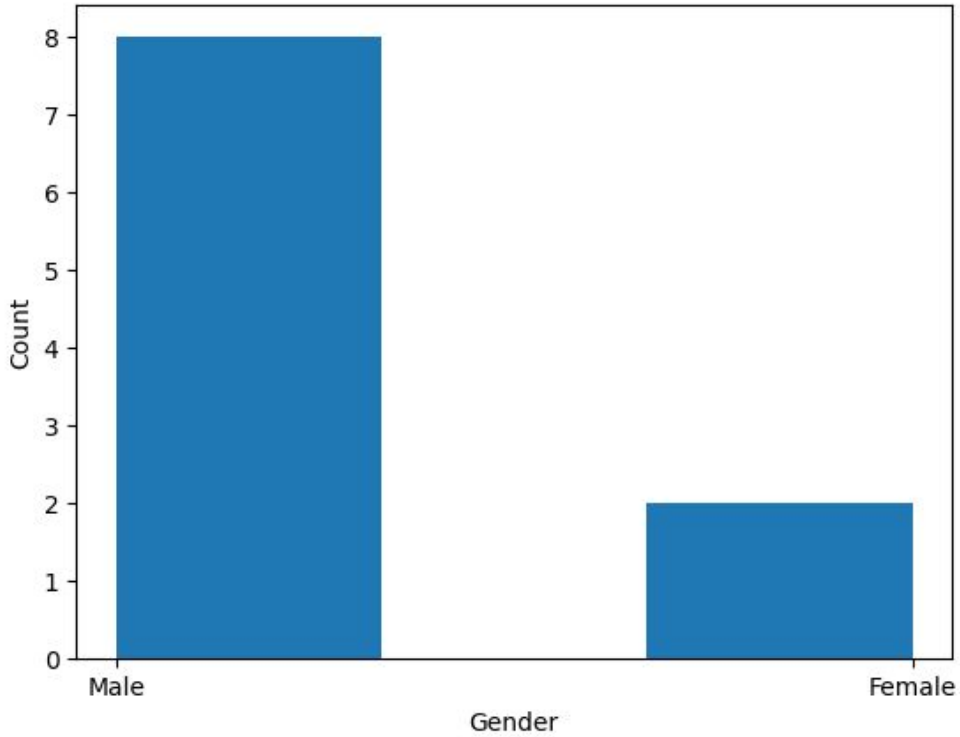
Demographics



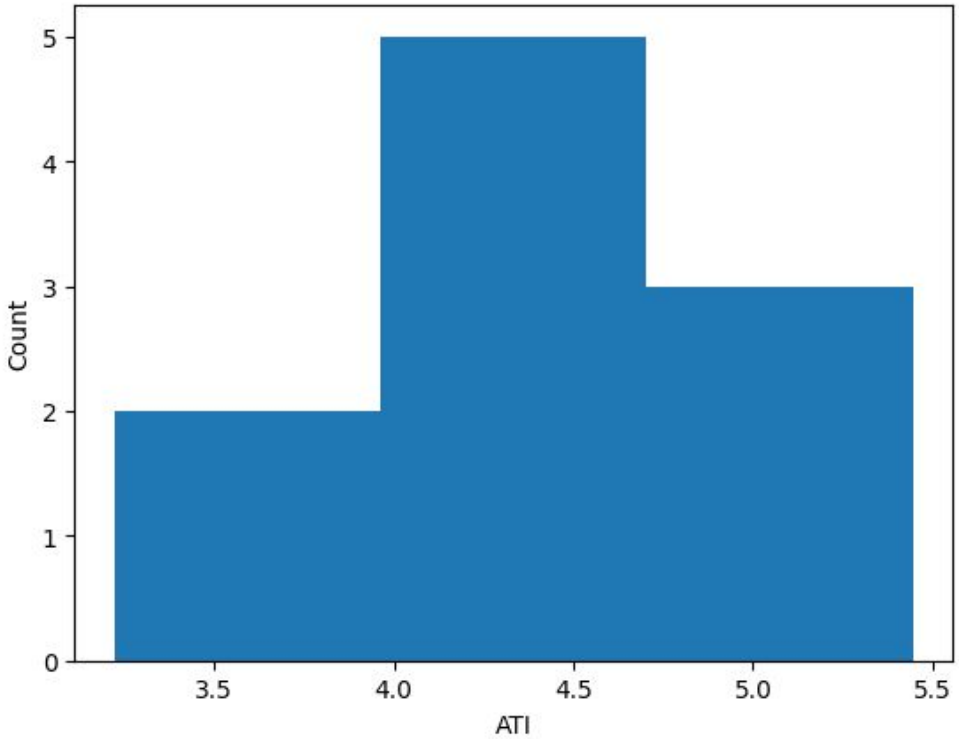
Demographics



Demographics

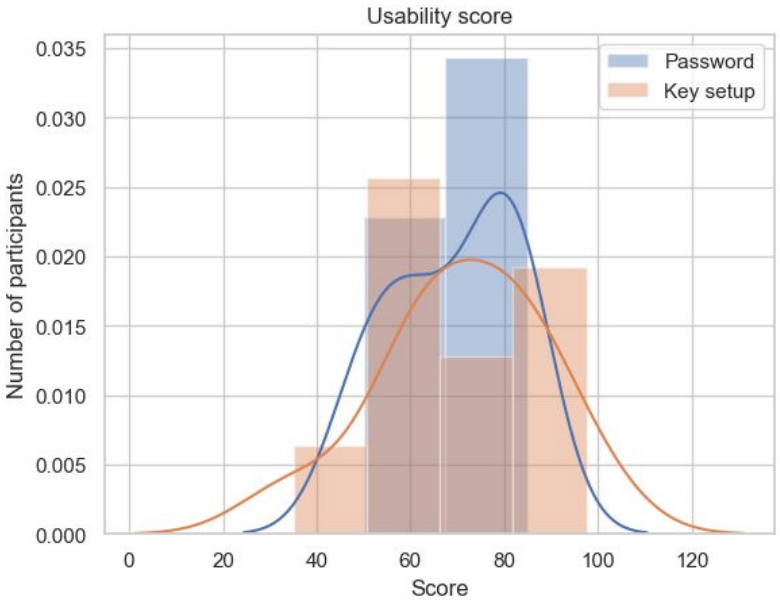


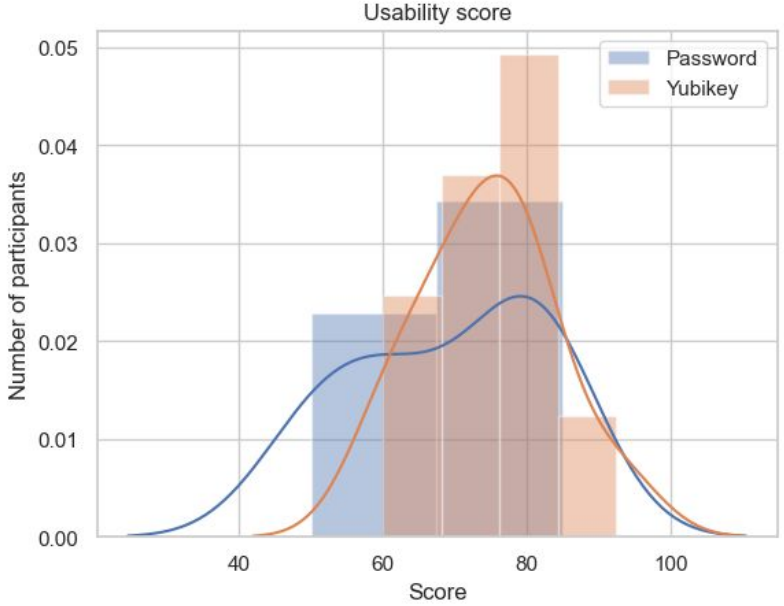
ATI



Average 4.3 meaning most participants are interested in technology

SUS


Usability score
Password
Key setup
Number of participants
0.035
0.030
0.025
0.020
0.015
0.010
0.005
0.000
0
20
40
60
80
100
120
Score


Usability score
Password
Yubikey
Number of participants
0.05
0.04
0.03
0.02
0.01
0.00
40
60
80
100
Score

Compared to paper

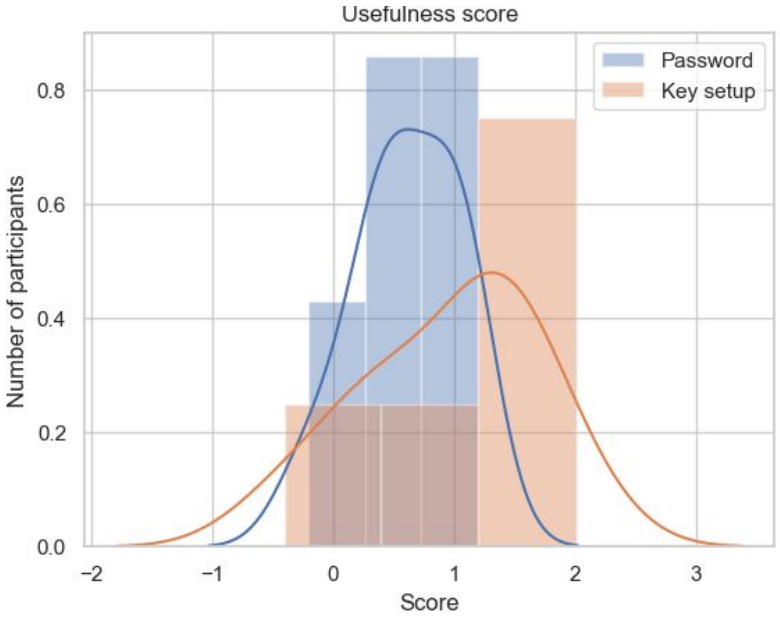
Paper:

"H1, an unpaired two-sample t-tests showed significant higher SUS scores in Group1FA (M = 81.74) than in GroupPass (M = 71.77); t(92) = 4.116, p < .001, Cohen’s d = .85"

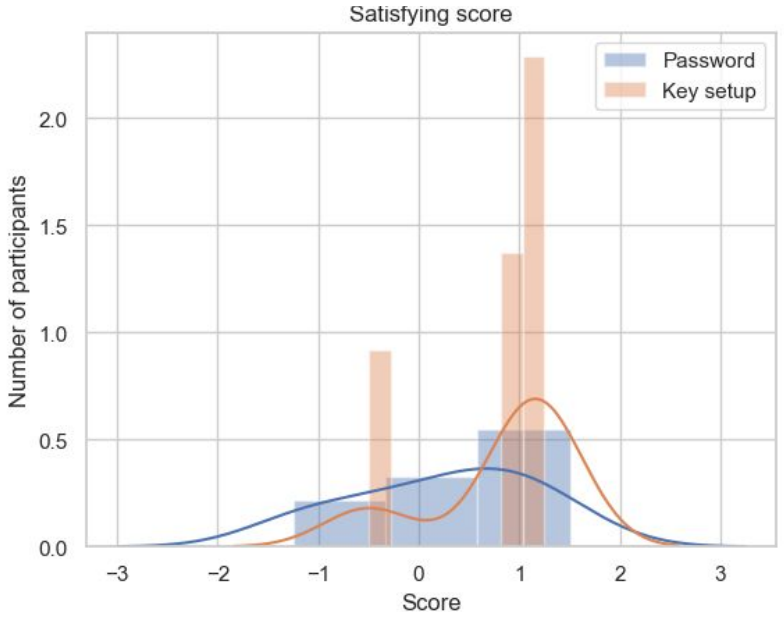
Our result:

- Password mean: 69.25
- Key: 71.0
- Mean difference: 1.75
- P-Value: 0.7819 >>> 0.001

Acceptance



P-Value: 0.2061 → insignificant



P-Value: 0.2059 → insignificant

Compared to paper

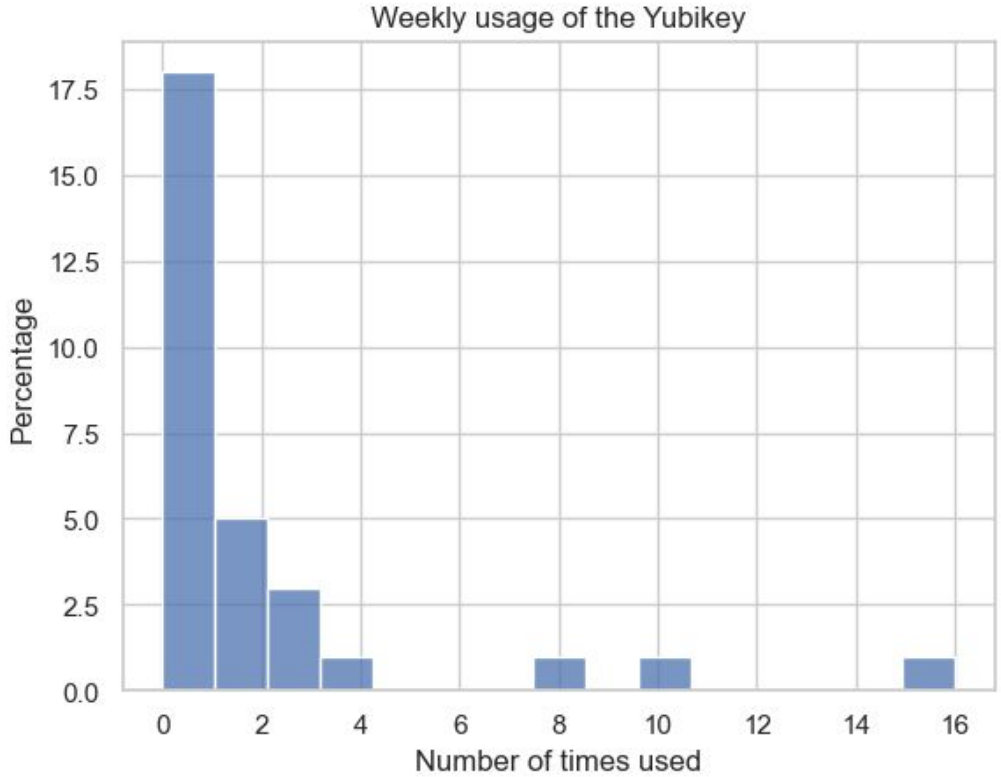
Paper:

"With respect to H2 an unpaired two-sample t-tests showed significant higher acceptance scores in Group- 1FA (M = 4.29) than in GroupPass (M = 3.41); t(92) = 6.522,p < .001, Cohen’s d = 1.35."

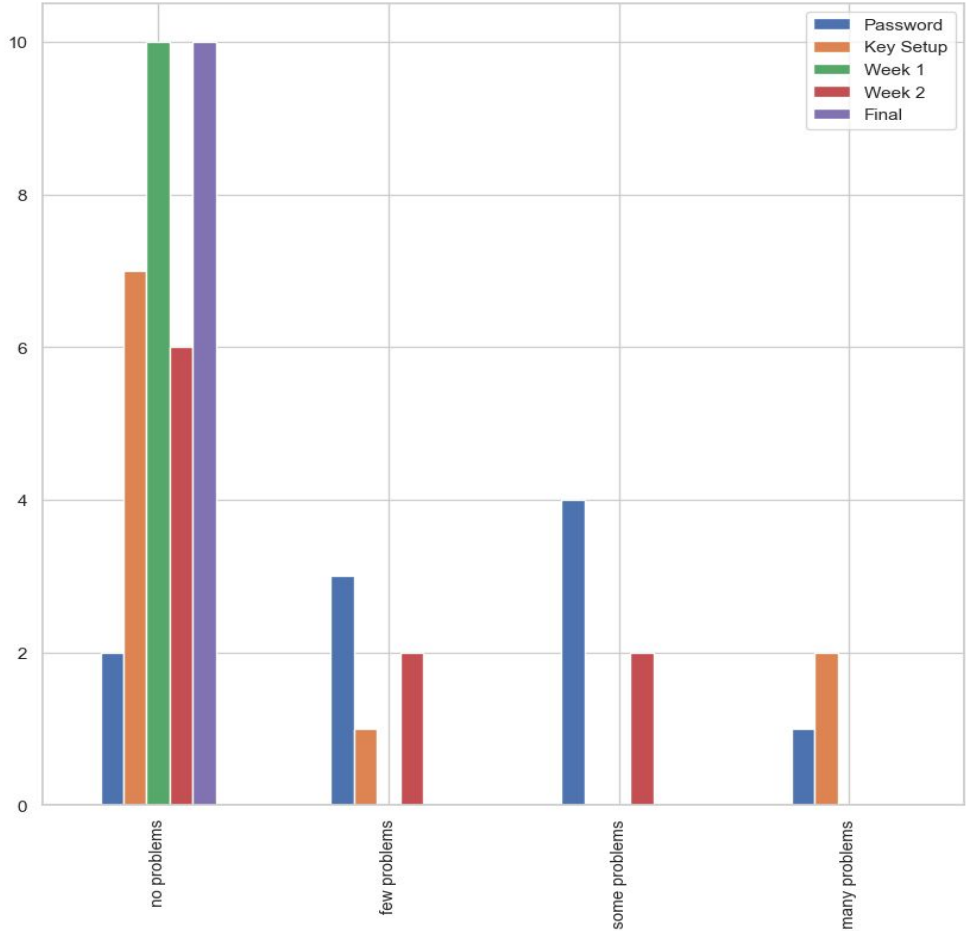
Our result:

- Password mean: 0,435
- Key: 0,8925
- Mean difference: 0.4575
- P-Value: 0.1724 >>> 0.001

Weekly usage



Weekly usage



Conclusion

- Neither H1 or H2 can be proven
- Some stuff can improved
- Still pending: more extensive analysis and open-ended question analysis

Thank you for attention