PANDERAM

Laboratory Study 2 PANDERAM Prototype Smartphone Security

Results AP 7.5 and 7.6

Susen Döbelt & Martin Köhler
Chemnitz, October 2023

secuvera
BSI-zertifizierter IT-Sicherheitsdienstleister und Prüfstelle

DAI-Labor
TU Berlin

Fraunhofer
SIT

ALLGEMEINE &
PSYCHOLOGIE
HUMAN FACTORS

INSTITUT
FÜR
TECHNIK
UND
JOURNALISMUS

Overview

Introduction: aims of the study, organization, procedure, and analysis

Results

- **Recruitment questionnaire**
- **Laboratory study 2**
- **Follow-up**

Summary

Outlook

Aim(s) of the study

- Investigation of the PANDERAM prototype with regard to potential improvements for **usability** and **user experience** as well as...

- **Fit** and **effectiveness** for different **behavioral stages**

Research Questions TUC/Investigation of:

Influence of the tailoring of the app (concruent/inconcruent) to the behavioral stage....

- ...on concern for privacy, perceived control over data, and effort expectation

Adapted behavioral stage instrument according to Bamberg (2013)

- Testing of model suitability for data protection context

Changes due to participation in laboratory experiment and confrontation with data protection issues

- Follow-up to capture (short-term) enduring changes in behavior and behavioral intention.

Organization 1/2

Timing

- Start conception (04/2023)
- Implementation and testing of questionnaires (06 & 07/2023)

- Start (recruitment): 06/27/2023
- Delivery of final prototype secuvera: 07/13/2023
- Start laboratory test: 07/19/2023
- End of laboratory test: 08/04/2023 (2.5 weeks test execution)
- Start Follow-Up: 07/29/2023
- Completion of follow-up: 08/14 & 16/2023 (in each case 10 days offset after trial implementation).
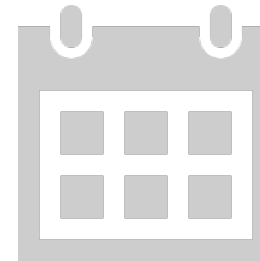
Organization 2/2

Recruitment

- Study call with link to recruitment questionnaire
- **Remuneration** laboratory test: 1 VP-h or 20€ in cash,
  Remuneration follow-up: 0.5 VP-h or 10€ (bank transfer or cash).
- **Dissemination:** via study participation distribution list TUC (2 reminders), AHF website,
  Notices in Wilhelm-Raabe-Str./TU Campus/Mensa/Library, chat groups & private environment.

A total of N = 82 individuals completed the recruitment questionnaire.

- Criteria for selection laboratory test:
  - Android operating system & >6.0 (n = 45, = 55% from recruitment questionnaire).
  - Behavioral stage assignment possible (n = 81) and balance of groups (1: VS 01 - 03 and 2: VS 04 - 05).

Ultimately, 23 individuals participated in the laboratory experiment. One person had to be excluded from the analysis due to technical difficulties during data acquisition **(n = 22).**

These n = 22 subjects also completed the follow-up questionnaire.

Organization 2/2

Recruitment

- Study call with link to recruitment questionnaire
- **Remuneration:** Remuneration for the laboratory test: 1 Remuneration follow-up: 0.5 subject hours or 10€ (bank
- **Dissemination:** via study participation distribution list T Notices in Wilhelm-Raabe-Str./TU Campus/Mensa/Lib

A total of _N_ = 82 individuals completed the recruitment

- Criteria for selection laboratory test:
  - Android operating system & >6.0 (_n_ = 45, = 54.9% from recr
  - Behavioral level assignment possible (_n_ = 81) and balance

Ultimately, 23 subjects participated in the laboratory experi analysis due to technical difficulties during data collection

These _N_ = 22 subjects also completed the follow-up que
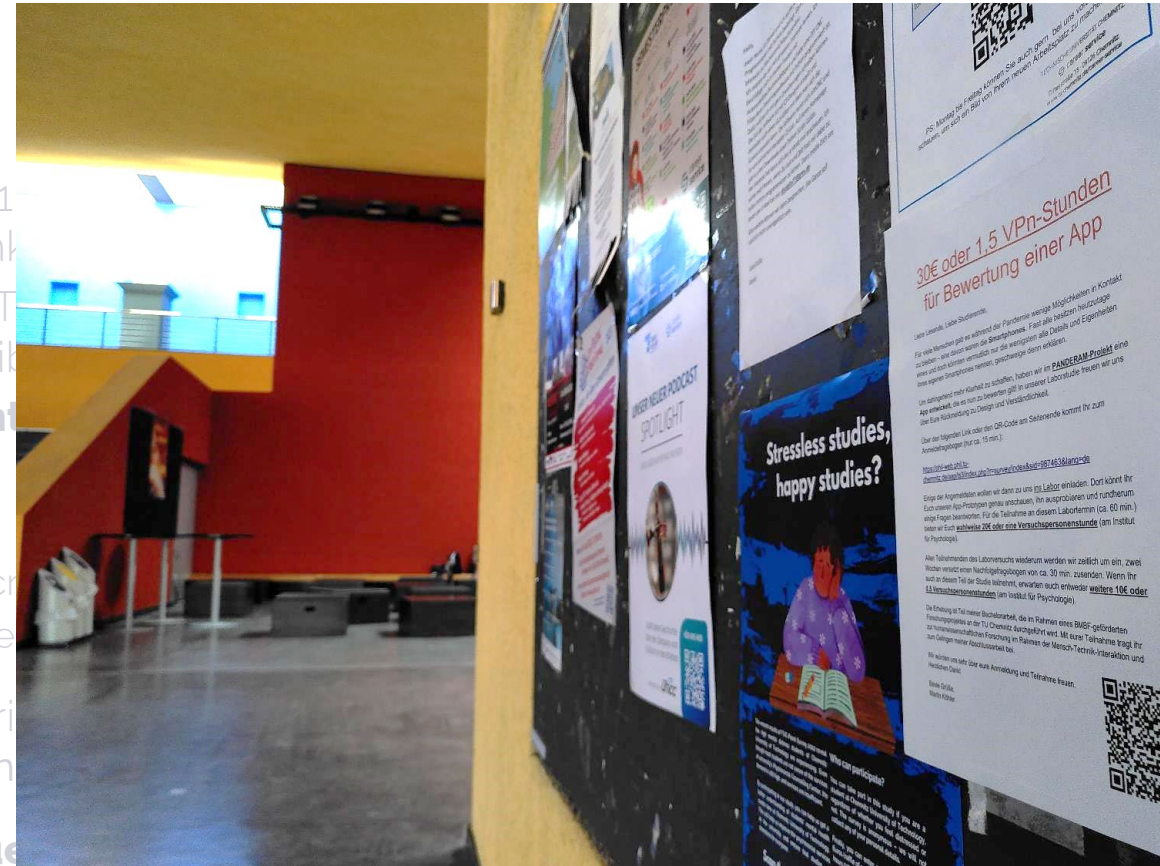
Procedure 1/8

Recruitment questionnaire

- **Welcome** and purpose of the registration questionnaire, privacy policy and consent form.
- **Demographic information** (age, gender, educational attainment, professional degree, current employment).
- Information on **smartphone usage** (operating system, version)
- Recording of the **behavioral stage** according to Bamberg (1a & 1b = Predecision, 2 = Preaction, 3 = Action, 4 = Postaction).
- Self-assessment of **technology affinity** and **smartphone competence**
- **Concern for privacy** (general) and specific to **smartphone use.**
- **Perceived control** when using the smartphone
- **Effort expectation** regarding the increase of smartphone security
- Forwarding to **separate contact deposit** and **remuneration request** (raffle or VP-h) indicating the subject code.
- **Adoption**

Completion of the **recruitment questionnaire: average time 12.74 min** (*SD* = 12.39; min = 4.20; max = 84.68).

Procedure 2/8

Before starting the examination

- **Rejection of** unsuitable persons from the recruitment questionnaire and **acceptance** via e-mail to selected pers
  - Content: Information about selection, request for **date selection** in the digital calendar (nuudel, Digitalcourage e.V.)
- For entry: **confirmation appointment,** sending **room information**
- Assignment of the experimental line and **assignment to the experimental condition** (prototype suitable or unsuitable for the behavioral stage).
- To the appointment: cleaning and ventilating the room, charging the test smartphone (Samsung Galaxy A33 5G, Android 12), resetting settings (lock screen & debugging if necessary), booting up test laptop and starting the questionnaire, getting participation information and privacy information ready, picking up the test subject.

Procedure 3/8

Before starting the examination

- **Welcome** and request to read and sign the **participation and privacy information.**

- Information on available **equipment**: laptop, keyboard and mouse for filling out the questionnaire, test smartphone as well as PANDERAM prototype installed on it.

- **Notes** from Trial Management:

    - Please do not delete any apps from the test smartphone and do not perform any updates

    - Follow instructions in the questionnaire and keep an eye on the timer during tasks

    - If questions arise, please contact the test management

    - Answer from the gut (there are no right or wrong answers)

Procedure 4/8

Investigation

Start of the experiment and self-assessment (pre-measurement)

- Enter subject code again
- **Effort expectation** regarding the increase of smartphone security
- **Perceived control** when using the smartphone
- **Privacy concern** specific to **smartphone use**
- Pre-measurement **of self-efficacy expectations regarding information security in** smartphone use

Procedure 5/8

Investigation

Task 0: **Getting to know** the PANDERAM prototype (10min time, please do not change behavioral stage setting)

Task 1 **Inform:** What are the security risks on the smartphone?

• Briefly describe the three most serious (5min time).

Task 2 **Action options User settings**: [...] First set up a lock screen with a password and then set an automatic screen lock. [...]

• What has changed about smartphone security by setting up the lock screen? Briefly describe the effect in your own words. (5min time)

Task 3 **Inform about options for action Device settings**: [...] Check the security status of the smartphone and pay attention to how long it has been since the last system software update. Only inform yourself about possible updates. Please do not perform any of them and do not download any updates. [...]

• When was the smartphone last updated? Describe in your own words what consequences the update status could have for the smartphone's security level (5min time)

Procedure 6/8

Investigation

Evaluation PANDERAM prototype
Quantitative evaluation by:

- SUS General Usability
- UEQ attractiveness, transparency, efficiency, predictability, stimulation, originality

Qualitative evaluation by:

- What aspects of the PANDERAM app did you particularly like? Name three aspects that are most important to you.
- What opportunities for improvement do you see for the PANDERAM app? Name three aspects that are most important to you.

School grade (1-6)

Procedure 7/8

Investigation

Post-measurement

- **Effort expectation** regarding the increase of smartphone security
- **Perceived control** when using the smartphone
- **Concern for privacy** specific to **smartphone use**
- Pre-measurement **of self-efficacy expectations regarding information security in** smartphone use

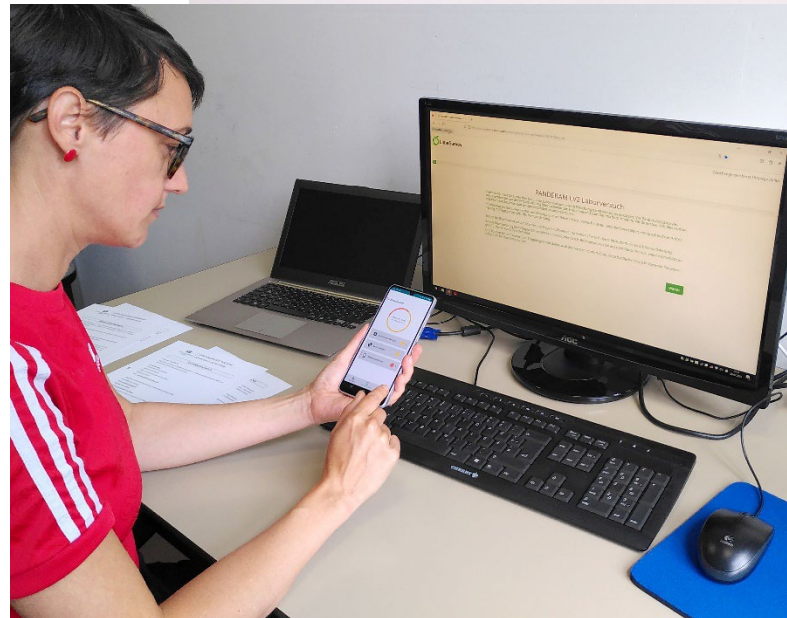**Goodbye** and **reminder of follow-up** questionnaire.

Payment of expense allowance or certification of subject hours.

Procedure 8/8

Investigation

Recruitment: demographics

Gender
■ male
■ female

32 %

68%

Age: *MW* = **23**.79 **years** (*SD* = 4.63; Min = 15; Max = 49).

**Gender**: 55 **female**, 26 male

Highest general educational qualification: 1.) **Abitur (university entrance qualification)** (77)
Highest vocational qualification: 1.) **Still in education** (28), 2. Bachelor (24), 3. none (10)

Current employment status: 1.) **Student** (67), 2.) employed (10),
3.) training/apprenticeship/(re)training

Course of study (*n* = 66): 1. **Psychology** (39), 2. Media Communication/Psychology (and Instructional Psychology) (6), 3. Computer Science (5).

Semester (*n* = 66): semester: *Mdn* = **4** (Min = 2; Max = 13)

Recruitment: smartphone use

Operating system

- Apple (iOS)
- Android

43%
56%
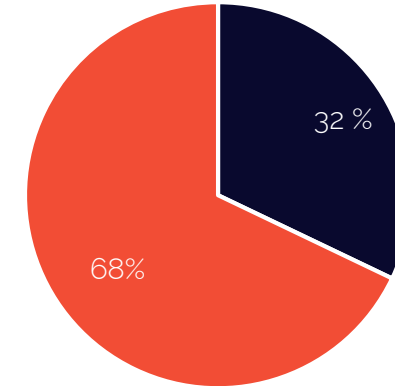
Use of a smartphone: 100% (81)

Smartphone **operating system**: 1.) **Android** 56% (45), 2.) iOS 43% (35)

Android operating system version (n = 45): 1.) **13** (16), 2.) 12 (12), 3.) 11 (7)
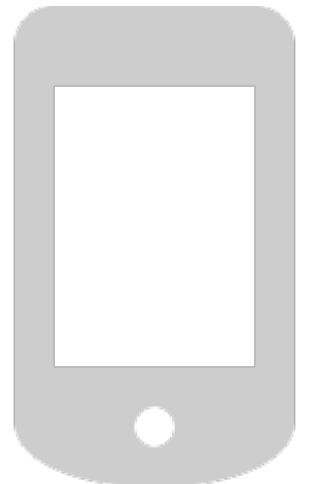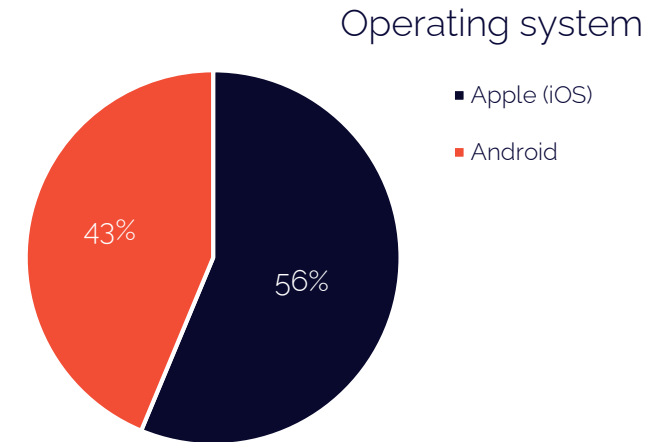
Market share operating systems smartphones worldwide, July `23 [01]:

Android 71%, iOS 28%

Market share of smartphone sales in Germany Q2 `23 [02]:

Android 70%, iOS 30%

The iOS operating system is **significantly more common** in our sample than in the national and global comparison.

[01] Statista (2023); [02] Statista (2023); https://iconmonstr.com/smartphone-1-png/

Recruitment: behavioral stages [03], [04].

Assignment to behavioral stages

Predecision: 21% (17)

Preaction: 32% (26)

Action: 22% (18)

Postaction: 24% (19)

N/A: 1% (1)

Behavioral stages
- Predecision
- Preaction
- Action
- Postaction
- N/A

1%
21%
32%
22%
24%
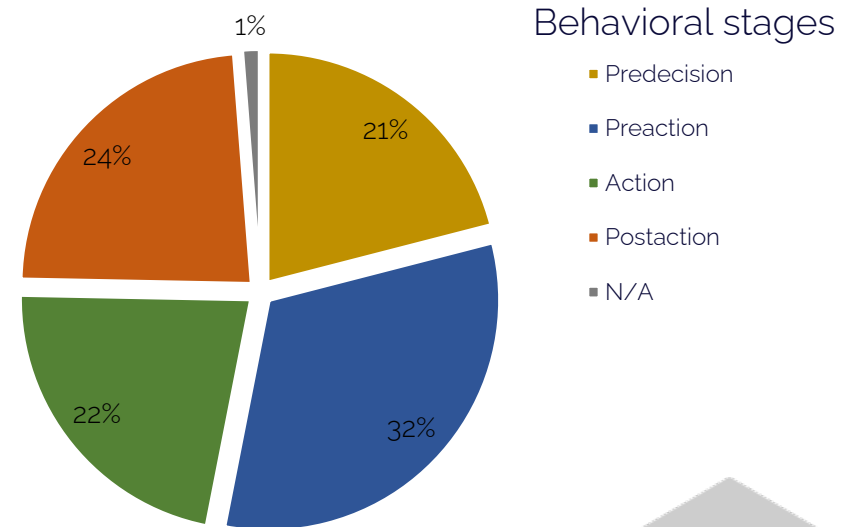
Most of the test persons *do not currently take any actions to protect their data. They are thinking about doing so, but do not yet know how.*

Only <u>one</u> respondent could not assign himself to any behavioral stages.

Compared to the previous PANDERAM study **the same ranking of** the behavioral stages and also **no** statistical **differences** in the occupation of the individual behavioral stages appeared.

[03] Bamberg, S. (2013a); [04] Bamberg, S. (2013b); https://iconmonstr.com/layer-21-png/

Recruitment: affinity for technology (ATI Scale; [05]).

Mean agreement: M = 3.82 („**somewhat agree**"; SD = .97; min = 1.78; max = 5.67, Mdn = 4.00).
Comparison with norm sample revealed **no** significant differences.

No group differences between behavioral stages.

Mean Affinity for Technology — Behavioral Stages (early, late); error bars: 95% CI

[05] Franke, T., Attig, C., & Wessel, D. (2019); https://iconmonstr.com/smartphone-15-png/
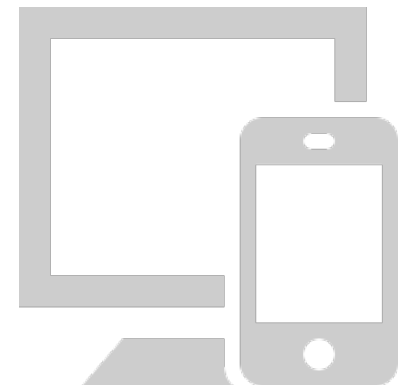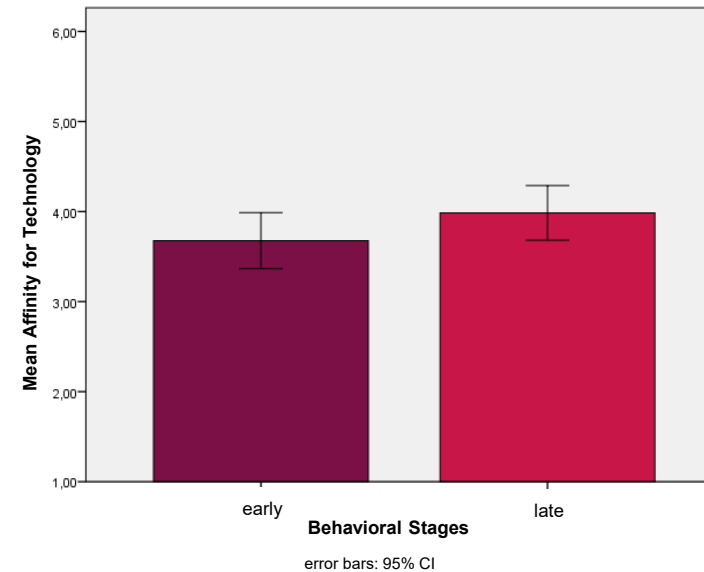
Recruitment: smartphone literacy (TAEG; [06]).

Mean agreement: M = 3.92 (= "***somewhat agree***"; SD = .6; Min = 2.75; Max = 5.00, Mdn = 4.00).

Comparison with first PANDERAM study (N = 99) revealed significant differences. Here, participants rate themselves as more competent on the smartphone than the comparison sample.

No group differences between behavioral stages.

error bars: 95% CI

[06] Karrer, K., Glaser, C., Clemens, C., & Bruder, C. (2009); https://iconmonstr.com/smartphone-1-png/
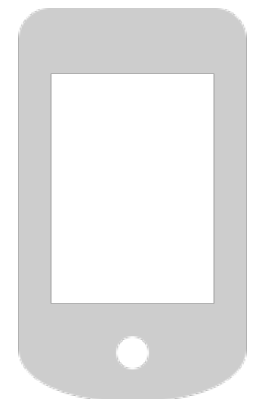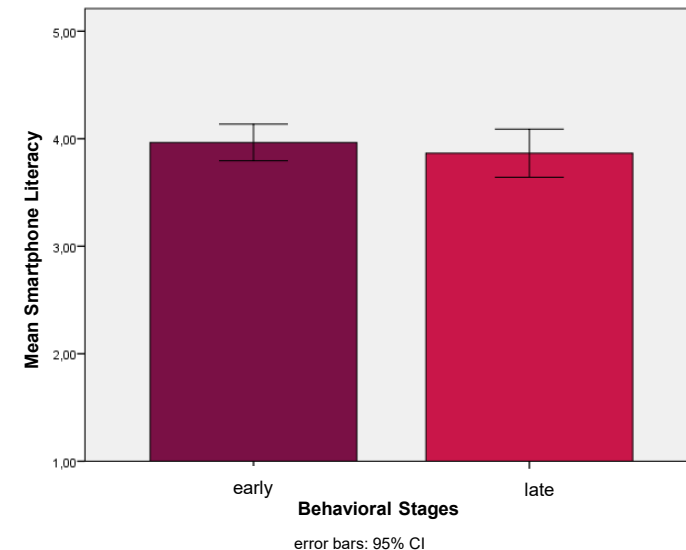
Recruitment: privacy concerns (IUIPC; [07]).

Mean Agreement Scale *Global Information Privacy Concerns*:
M =4.26 (= "***somewhat agree***"; *SD* = 1.09; Min = 1.83; Max = 7.00)

Comparison with AndProtect study (*N* = 213) revealed significant differences . Here, participants rate their concern for privacy lower.

Differences between behavioral stages given.

Participants in the late behavioral stages rate their concern for privacy higher than those in early behavioral stages.

error bars: 95% CI

[07] Malhotra, N. K., Kim; S. S., & Agarwal, J. (2004); https://iconmonstr.com/cloud-31-png/
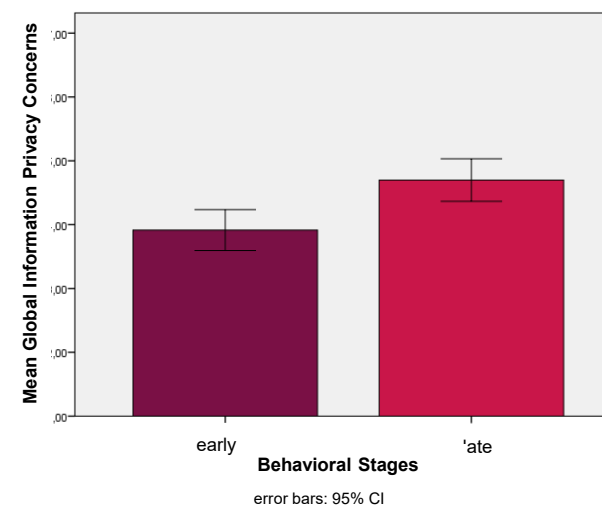
Recruitment: concern about smartphone security
(PCON; [08])

Mean agreement: *M* = 4.41 (= "***undecided***"; *SD* = 1.21; Min = 2.00; Max = 7.00)

Comparison with sample means (*N* = 222) from Yun et al. (2013) [09] revealed significant differences. Here, participants rate their concerns lower.

No group differences between behavioral stages.
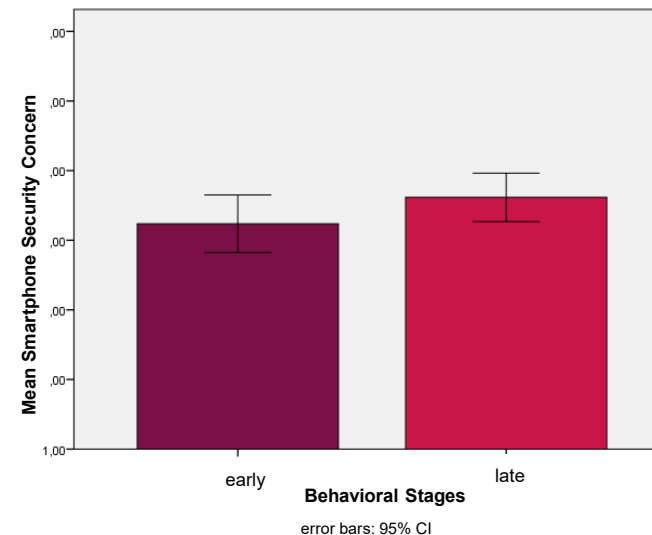
Recruitment: effort expectation (UTAUT 2; [10]).

Mean agreement: *M* = 3.95 (= "***somewhat agree***"; *SD* = .99; Min = 2.00; Max = 6.50).

Comparison with sample means (*N* = 198*)* from Mangiò et al. (2020) [11] revealed significant differences. Participants expect more effort (agree to statements less) than comparison sample.

No group differences between behavioral stages.

[10] Venkatesh, V., Thong, James Y. L., Xu, X. (2012); [11] Mangiò, F., Andreini, D. & Pedeliento, G. (2020); https://iconmonstr.com/construction-12-png/
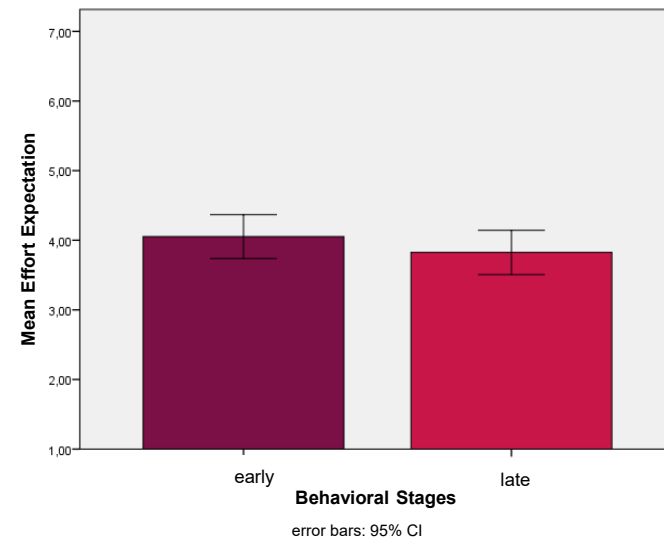
Recruitment: perceived control (PCTL; [08])

Mean agreement: M = 3.51 (= "***rather no control***"; SD = .96; Min = 1.00; Max = 5.50).

Comparison with sample means (N =128) from Zhang & Xu (2016) [12] revealed significant differences. Here, participants perceive less control than comparison sample.
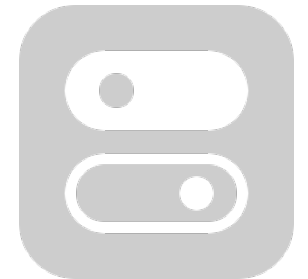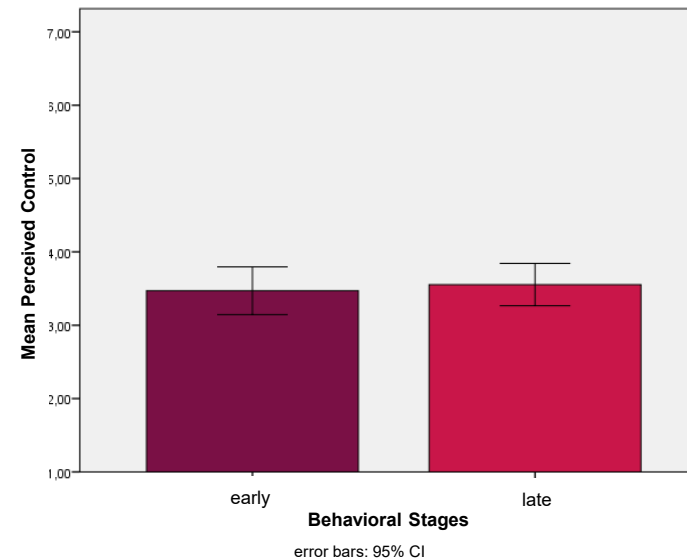
No group differences between behavioral stages.

Mean Perceived Control vs. Behavioral Stages (early, late), error bars: 95% CI

[08] Xu, H. (2007); [12] Zhang, B. & Xu, H. (2016); https://iconmonstr.com/control-panel-3-png/

Recruitment: questions & hypothesis

Q2: Are there a-priori differences between behavioral stages in smartphone users' concern for their smartphone security, effort expectations, and perceived control over their data?

H2a: There are a-priori differences in smartphone users' **concern for** their smartphone security depending on behavioral stage.

Result: hypothesis (almost) confirmed

error bars: 95% CI

Recruitment: questions & hypothesis

Q2: Are there a-priori differences between behavioral stages in smartphone users' concern for their smartphone security, effort expectations, and perceived control over their data?

H2b: There are a-priori differences in smartphone users' **effort expectations to** increase their smartphone security depending on <u>behavioral stage</u>.

Result: hypothesis not confirmed

error bars: 95% CI

Recruitment: questions & hypothesis
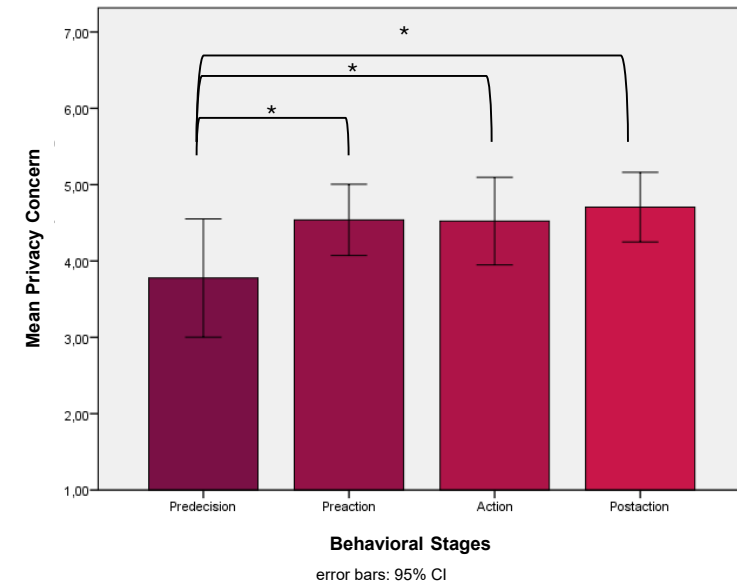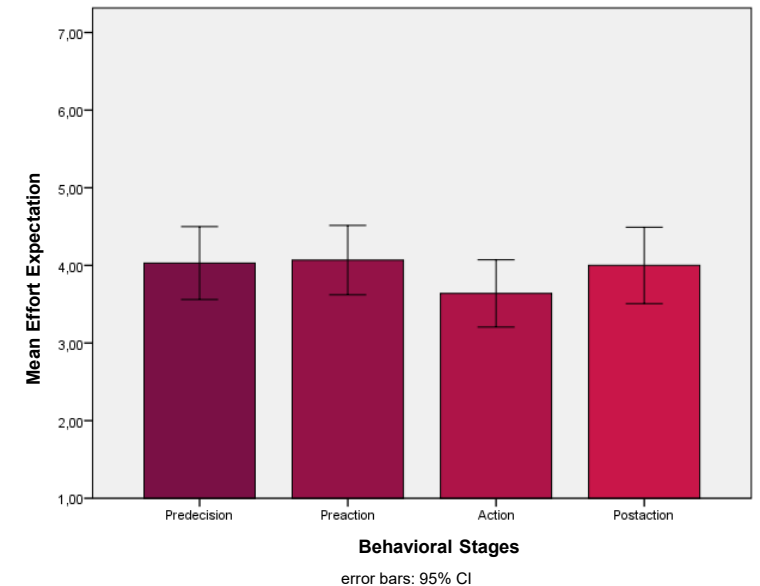
Q2: Are there a-priori differences between behavioral stages in smartphone users' concern for their smartphone security, effort expectations, and perceived control over their data?

H2c: There are a-priori differences in smartphone users' **perceived control** over their data depending on <u>behavioral</u> stages.

Result: Hypothesis not confirmed
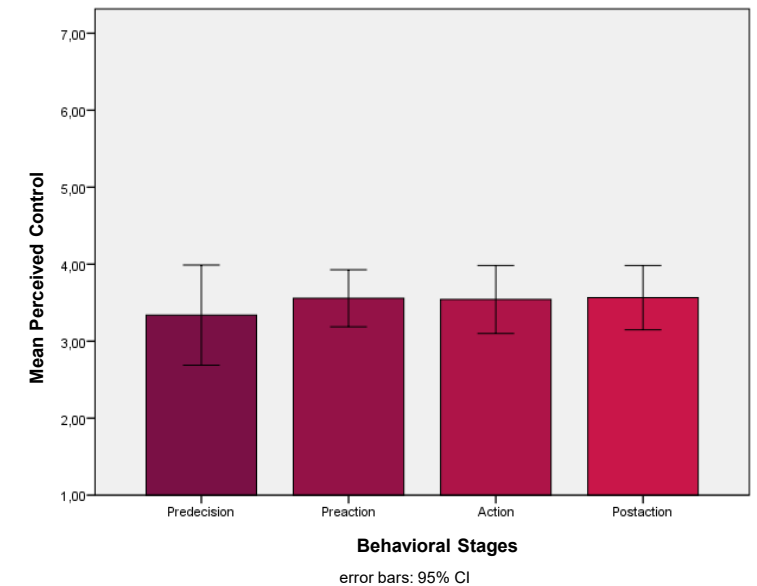
Recruitment: interim conclusion

A **(student) sample** was approached for the 2nd laboratory test, which is similar to laboratory test 1. This sample estimates itself to be **rather unconcerned about** privacy protection as **well as competent in the use of smartphones**. Participants in the recruitment indicate that they **perceive less control over their data** and **expect more effort to** increase their smartphone security than comparable samples.

Individuals of different **behavioral** stages **differ** statistically only **with respect to** their stated **concern**:

On the one hand, in the **general concern for privacy** between the earlier and later behavioral stages groups.

On the other hand (only marginally significant) **in concern about their smartphone safety** between individuals in the earliest behavioral stages ("*Predecision*") and all others.

Laboratory: demography (*n* = 22)

Gender
- male
- female

32%

68%

Age: *M* = 24.82 (*SD* = 4.44; Min = 19; Max = 36).

Gender: 15 female, 7 male

Highest educational attainment: 1.) Abitur (university entrance qualification)

(21), 2.) secondary school certificate (Realschulabschluss) (1)

Highest vocational qualification: 1.) None (yet) (9), 2.) Bachelor's degree (6), completed,

3.) Training/vocational school (6)

Current employment status: 1.) student (16), 2.) employed (5), 3.) in training (1)

Degree programs (*n* = 16)**:** 1. psychology (9), 2. computer science (4), 3. media comm./psych. (2),

4. teaching (1).

Semester (*n* = 16)**:** *MW* = 3.63 (*SD* = 2.10; Min = 2; Max = 8)

Behavioral stages: BS 1 = 4, BS 2 = 6, BS 3 = 4, BS 4 = 8

Concruent (12) and inconcruent (10) variant
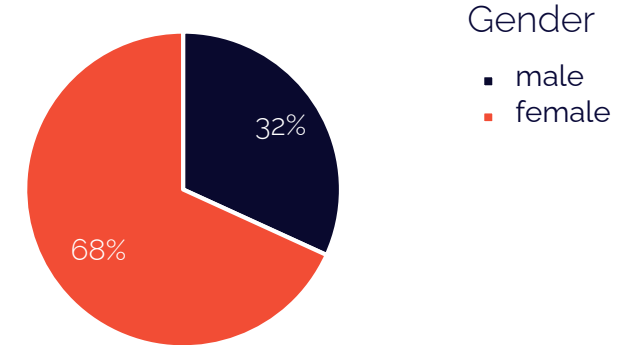
Laboratory: affinity for technology (ATI Scale; [05])

Mean agreement: *M* = 3.91 (="***somewhat agree***"; *SD* = .95; min = 1.78; max = 5.00, *Mdn* = 4.00).

Comparison with norm sample (Wilcoxon test) (*N* = 1567; [5]) revealed **no** significant differences.

No group differences between behavioral stages.

error bars: 95% CI

[05] Franke, T., Attig, C., & Wessel, D. (2019); https://iconmonstr.com/smartphone-15-png/
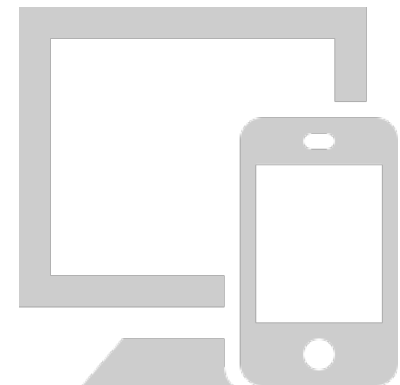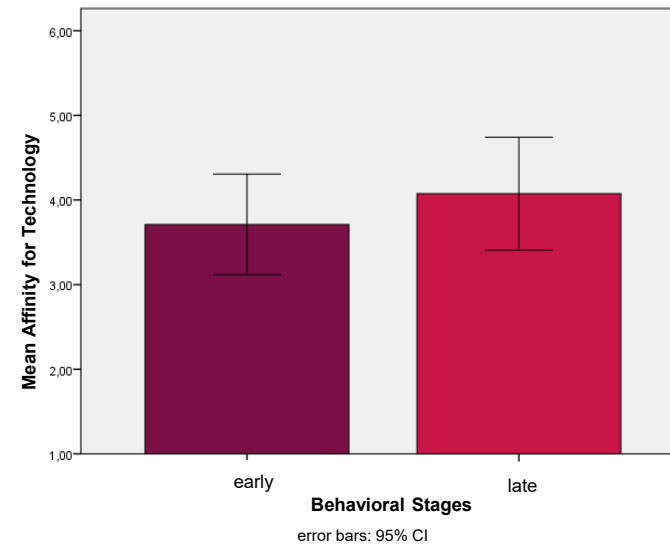
Laboratory: smartphone literacy (TAEG; [o6]).

Mean agreement: *MW* = 3.76 (= "***somewhat agree***"; *SD* = .60; Min = 2.75; Max = 4.75)

Comparison with sample means (*N* = 460) from Karrer et al. (2009) revealed significant differences. Our participants rate themselves as more competent.

No group differences between behavioral stages.

error bars: 95% CI

[11] Karrer, K., Glaser, C., Clemens, C., & Bruder, C. (2009); https://iconmonstr.com/smartphone-1-png/
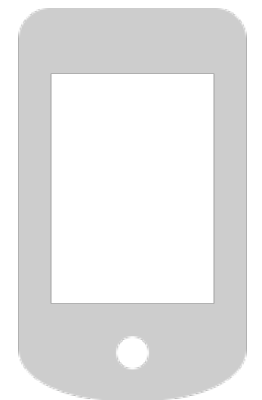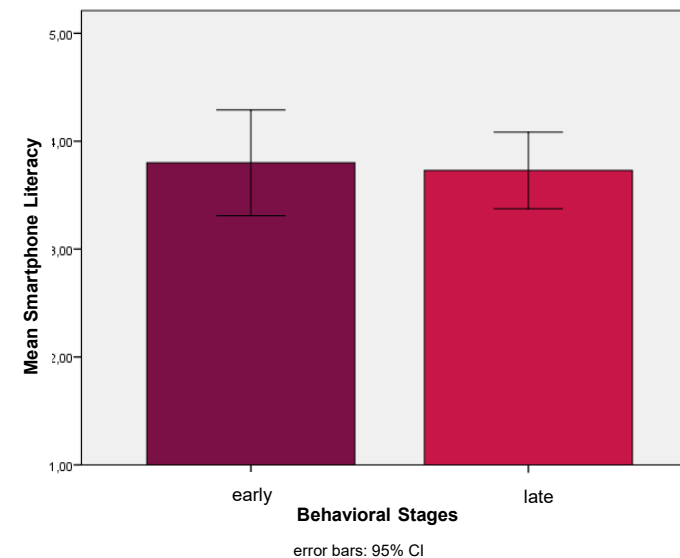
Laboratory: privacy concerns (IUIPC; [07]).

Mean agreement: M = 4.70 (= "***somewhat agree***"; SD = .95; Min = 3.17; Max = 6.67).

Comparison with AndProtect study means (N = 213) revealed significant differences. Our participants rate themselves as less concerned.

No group differences between behavioral stages.
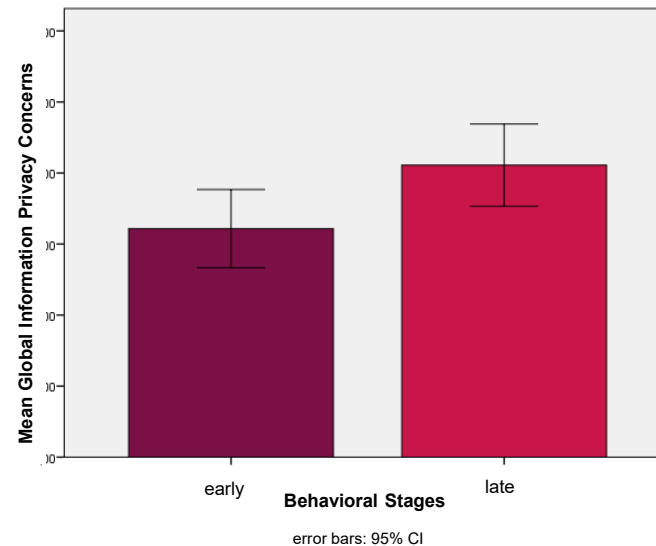
error bars: 95% CI

[07] Malhotra, N. K., Kim; S. S., & Agarwal, J. (2004); https://iconmonstr.com/cloud-31-png/

Laboratory: questions & hypothesis

Q1: What impact does have the tailoring of the PANDERAM app variant to the behavioral stage on changes in smartphone users' concern about their smartphone security, perceived control over their data, and effort expectation from the intervention?

H1a: When participants in the earlier behavioral stages are presented with the appropriate **app variant**, it increases their **concern for** their smartphone safety more than when they are presented with the inappropriate app variant.

Result: Hypothesis (partly) confirmed
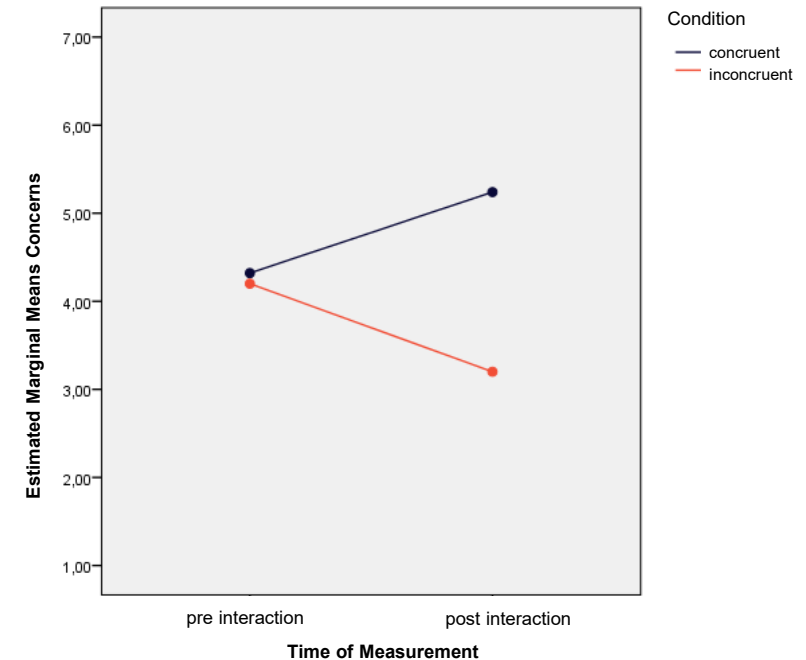
34

Laboratory: questions & hypothesis

Q1: What impact does have the tailoring of the PANDERAM app variant to the behavioral stage on changes in smartphone users' concern about their smartphone security, perceived control over their data, and effort expectation from the intervention?

H1b: When participants in the later behavioral stage are presented with the appropriate **app variant**, this reduces their **concern about** their smartphone safety more than when they are presented with the inappropriate app variant.

Result: Hypothesis not confirmed
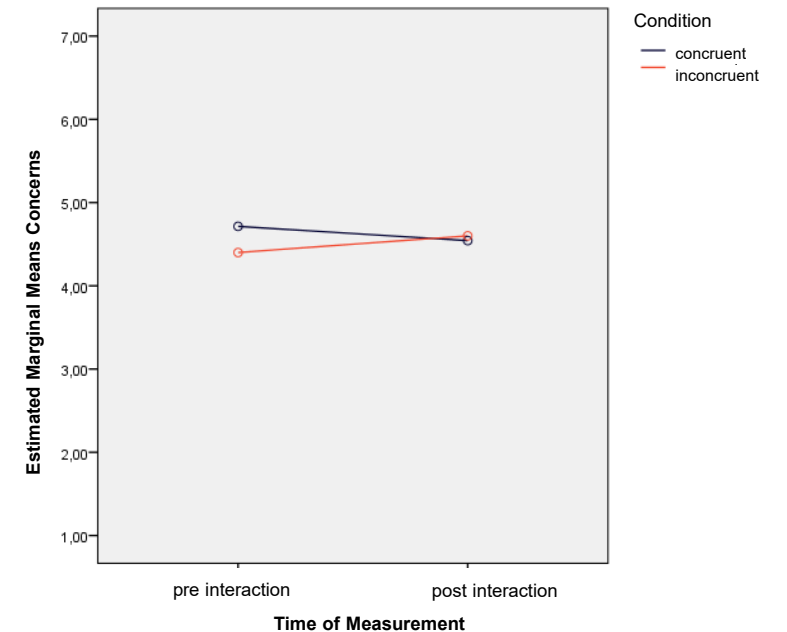
Laboratory: questions & hypothesis

Q1: What impact does have the tailoring of the PANDERAM app variant to the behavioral stage on changes in smartphone users' concern about their smartphone security, perceived control over their data, and effort expectation from the intervention?

H1c: If participants are presented with the appropriate **app variant**, this reduces their **effort expectation** to increase their smartphone security more than if they are presented with the inappropriate app variant.

Result: hypothesis not confirmed, but:

Main effect for time
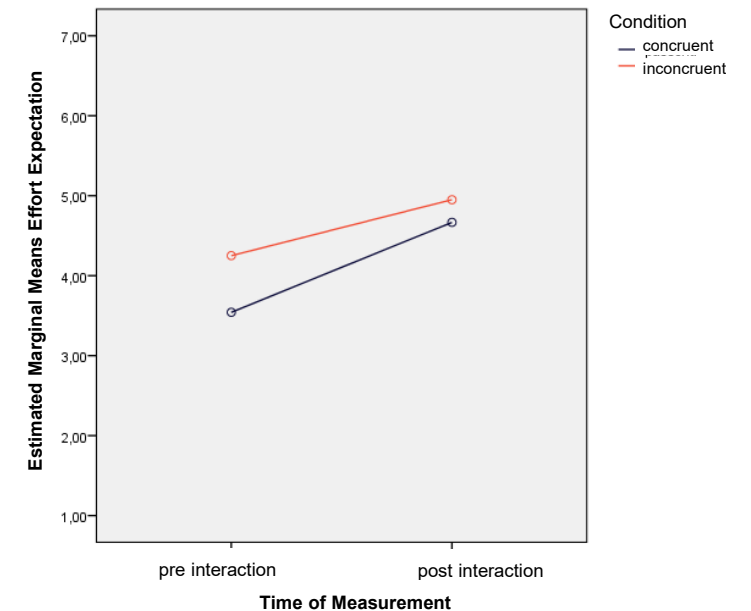
Laboratory: questions & hypothesis

Q1: What impact does have the tailoring of the PANDERAM app variant to the behavioral stage on changes in smartphone users' concern about their smartphone security, perceived control over their data, and effort expectation from the intervention?

H1d: Presenting participants with the appropriate **app variant** increases their **perceived control** over their data more than presenting them with the inappropriate app variant.

Result: hypothesis not confirmed, but:

Main effect time
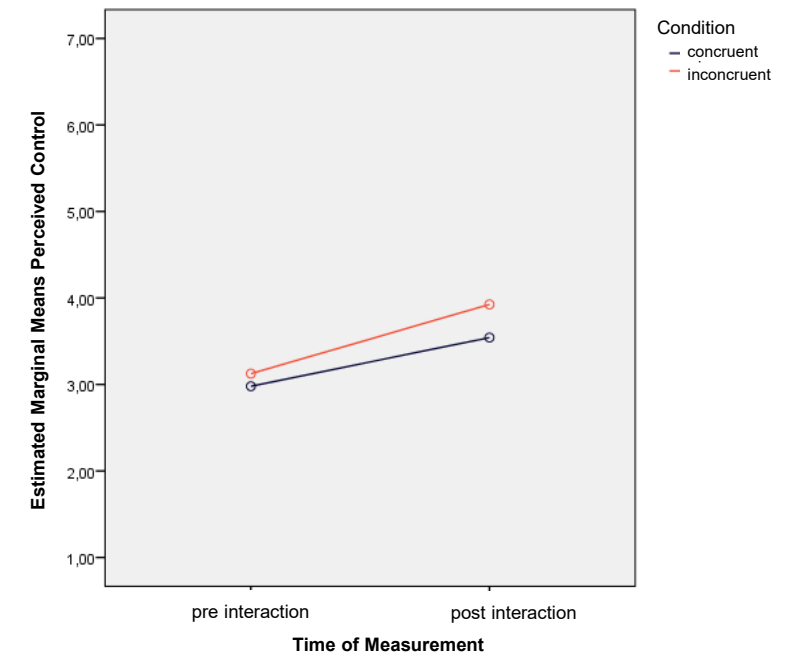
Laboratory: task 1 (qualitative)

What are the security risks on the smartphone?
Briefly describe three of the most serious ones.

There were **76 statements** classified into 8 categories, most frequently:

1.) Freeletics app (data transfer, third-party providers, use of outdated methods, connections) (26%)

2.) Necessary software update of the smartphone (24%)

3.) Unlock screen without password or fingerprint (20%)

Security risks mentioned

Necessary software update smartphone	24%
Unlock screen without password/fingerprint	20%
Automatic screen lock (deactivated or too long)	11%
Risk Freeletics app (data transfer, third-party providers, use of outdated processes, connection, etc.)	26%
Location transmission	7%
USB debugging activated	5%
Data readout by unauthorized persons	4%
Misc.	4%

0%   10%   20%   30%

Relative response frequency

38

Laboratory: Task 2 (qualitative)

What has changed about smartphone security by setting up the lock screen? Briefly describe the effect in your own words.

There were **48 statements** classified into 6 categories, most frequently:

1.) Change in security level PANDERAM app (23%)

2.) Increased protection of data (no access) (23%)

3.) Locking the screen by itself (23%)

4.) Password is required (23%)

Named changes by lock screen setup

Category	%
Changing the security level in the Panderam app	23%
Increased data protection (no access to smartphone data)	23%
Automatic screen lock (timeout)	23%
Password is required	23%
Misc.	6%
No Answer	2%

0%   10%   20%   30%

Relative response frequency
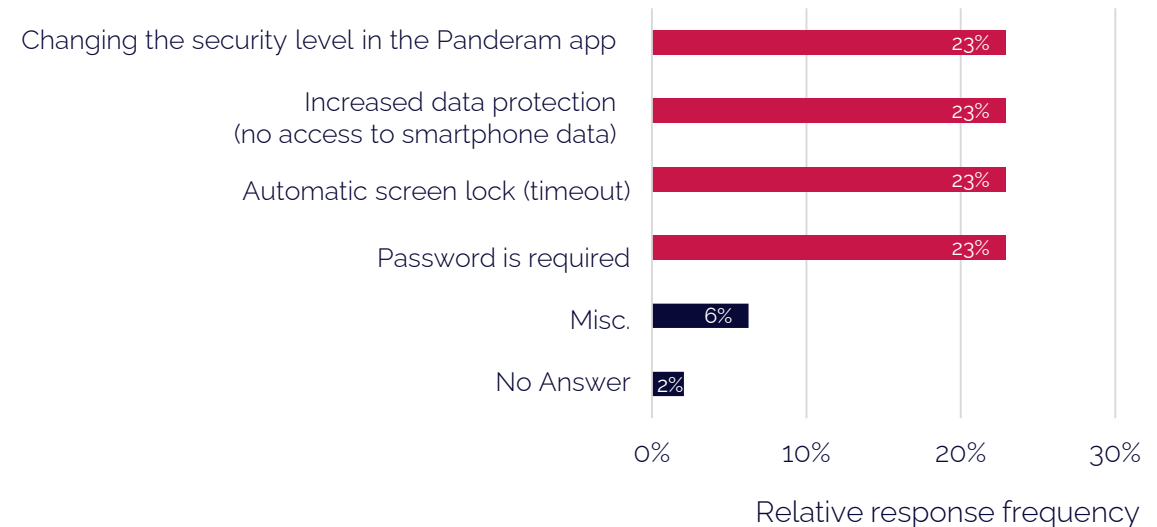
Laboratory: task 3 (qualitative)

When was the smartphone last updated? Describe in your own words what consequences the update status might have for the smartphone's security level.

There were **48 statements** classified into 5 categories, mos
frequently:

1.) More security holes & vulnerabilities (46%),

2.) Last updated on 01.09.22 (33%), and

3.) Last updated on 30.06.23 (17%)

Update time and named effects

Last update: 06/30/23	17%
Last update: 01/09/22	33%
More security gaps and vulnerabilities without the latest update (access to data by external attackers, viruses)	46%
Pressure to buy new devices for current safety standard	2%
Misc.	2%

0%   10%   20%   30%   40%   50%

Relative response frequency

40
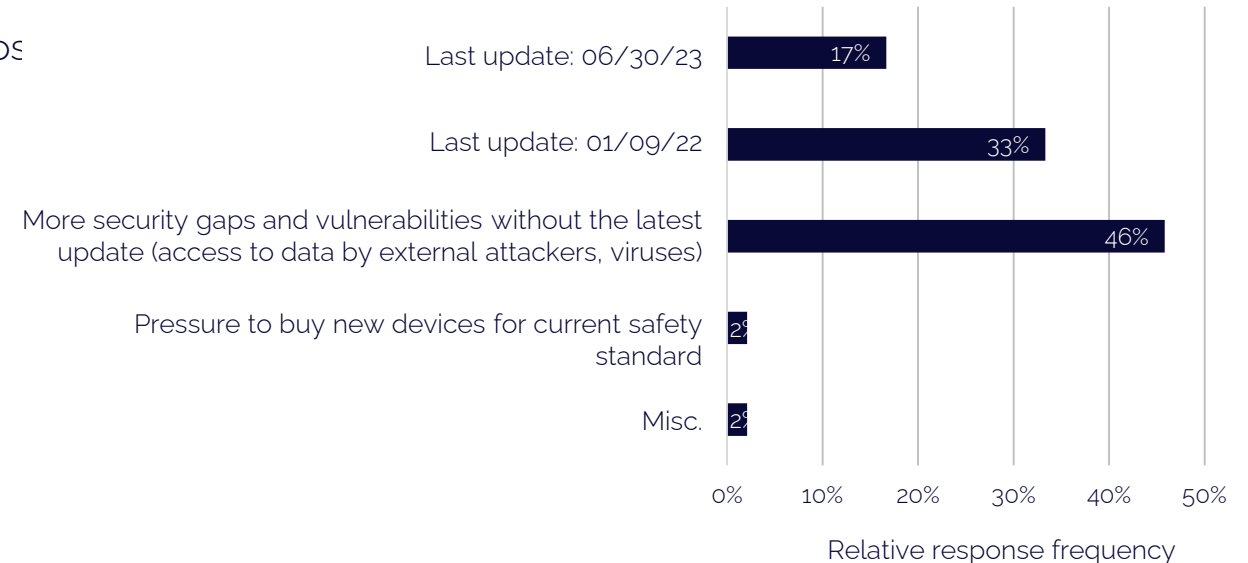
Laboratory: usability 1/2

System Usability Scale (SUS; [14])

Mean SUS score (n = 21): *MW* = 77.50 (*SD* = 9.59, Min = 52.50, Max = 90.00) of the PANDERAM app, corresponding to [14] **grade B+**.

Two-sided one-sample *t-test*: no difference t(20) = -1.20, p = .256, d = -.26 from laboratory test 1.
One-sided testing against benchmark [15]: SUS score in the middle "*good range*": significant difference t(20) = 2.34, p = .015, d = .51 to score 72.6 = lower limit of B-

The evaluation of the PANDERAM prototype with regard to **usability turns out to be good** overall and is **comparable to laboratory test 1**.

[14] Brooke, J. (1996); [15] Lewis & Sauro (2018); https://iconmonstr.com/check-mark-16-png/
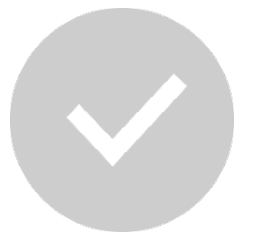
Laboratory: usability 2/2

System Usability Scale (SUS; [14])

Comparison between matching and mismatching condition

Descriptive statistics

Concruent condition: n = 11, M = 79.32, SD = 2.61
Inconcruent condition: n = 10, M = 75.50, SD = 10.59

No difference revealed between the app variants.

error bars: 95% CI

[14] Brooke, J. (1996); https://iconmonstr.com/check-mark-16-png/
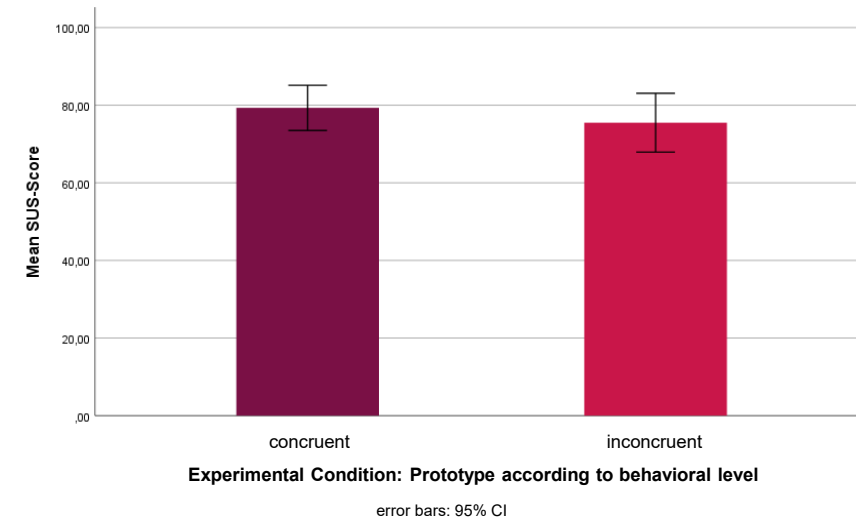
Laboratory: user experience 1/3

UEQ [16]

Scale	n	MW	SD	Min	Max	Rating
Attractiveness	21	1.74	.79	.50	3.00	good (to above average)
Transparency	21	2.07	.82	.25	3.00	excellent (to good)
Predictability	21	2.00	.55	1.25	3.00	excellent
Stimulation	21	1.61	.94	-.25	3.00	excellent (to above average)
Originality	21	1.08	1.01	-.33	2.33	good (to above average)
Efficiency	21	1.87	.84	.25	3.00	excellent (to good)
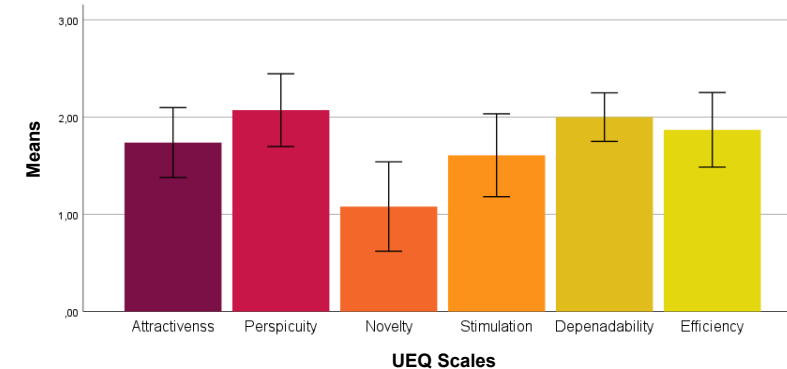
error bars: 95% CI

On average, UX is predominantly rated as "excellent". There are greater deviations in the assessment of the stimulation dimension. The attractiveness and originality of the app could be improved the most in the opinion of the participants.

[16] Laugwitz et al. (2008); https://iconmonstr.com/smiley-1-svg/

Laboratory: user experience 2/3

UEQ [16]

	Laboratory test 1					Laboratory test 2					
Scale	n	MW	SD	Min	Max	n	MW	SD	Min	Max	Difference (two-tailed)
Attractiveness	38	1.49	.72	-.83	3.00	21	1.74	.79	.50	3.00	no
Transparency	38	1.88	1.01	-2.25	3.00	21	2.07	.82	.25	3.00	no
Predictability	38	1.80	1.80	.50	2.75	21	2.00	.55	1.25	3.00	no
Stimulation	38	1.13	.77	2.50	3.25	21	1.61	.94	-.25	3.00	yes (LT 2 better)
Originality	38	1.07	.89	1.50	2.75	21	1.08	1.01	-.33	2.33	no
Efficiency	38	1.72	.65	-.25	3.00	21	1.87	.84	.25	3.00	no

Only the **stimulation** dimension was **rated better in laboratory test 2. The ratings on the other dimensions did not differ compared to laboratory test 1.**

[16] Laugwitz et al. (2008); https://iconmonstr.com/smiley-1-svg/

Laboratory: user experience 3/3

UEQ [16]

Comparison between concruent and inconcruent condition

All scales were rated better in the matching prototype variant than in the mismatching one. However, statistically no effect can be found for the experimental conditions.

Only in the **Novelty scale does the** difference become marginally significant:

$Mdn_{concruent}$ = 1.66 ; $Mdn_{inconcruent}$ = 0.55

Experimental Condition: **Prototype according to behavioral level**

error bars: 95% CI

UEQ: Novelty

Experimental Condition

error bars: 95% CI

[16] Laugwitz et al. (2008); https://iconmonstr.com/smiley-1-svg/
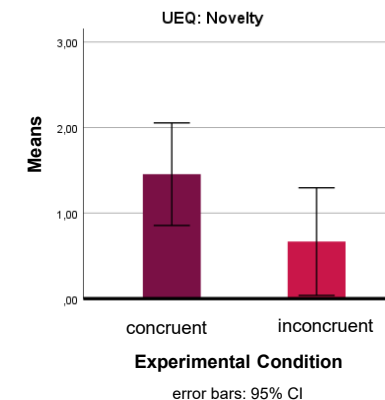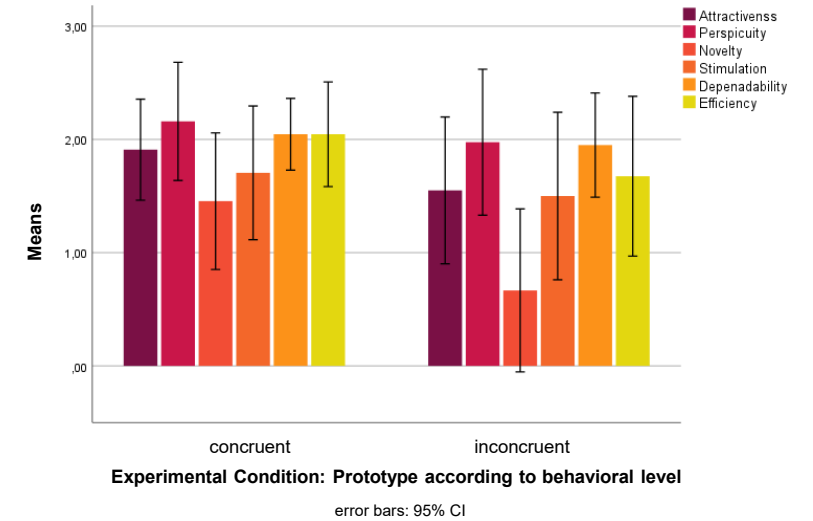
Laboratory: self-efficacy expectation 1/10

SWE [17]

General self-efficacy expectation according (SWE) to Schwarzer and Jerusalem (10 items on 4-point response scale ("*not true, hardly true, rather true, true exactly*")

 e.g., "*When resistance to protecting my privacy arises, I find ways to assert myself.*"

Measurement	n	M	SD	MIN	MAX
Pre-measurement	22	2.24	0.39	1.50	2.80
Post-measurement	22	2.58	0,49	1.40	3.70

The data before and after trying the prototype differ highly significantly.

The self-efficacy of the test subjects was significantly improved by the interaction with the prototype.

error bars: 95% CI

[17] Schwarzer, R., & Jerusalem, M. (2002); https://iconmonstr.com/wrench-24-png/
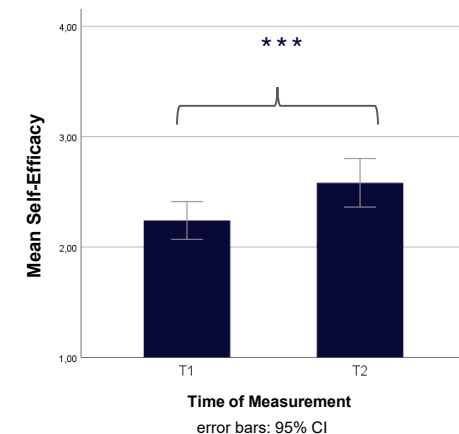
Laboratory: self-efficacy expectation 2/10

SWE [17]

Main effect time but not for experimental condition and not for the interaction time x experimental condition.

The subjects' self-efficacy was significantly improved by their interaction with the prototype. **It did not matter which prototype variant (suitable/incompatible) the test subjects tried out.**

Condition
- concruent
- inconcruent

Mean Agreement Self-Efficacy

Time of Measurement
error bars: 95% CI

[17] Schwarzer, R., & Jerusalem, M. (2002); https://iconmonstr.com/wrench-24-png/
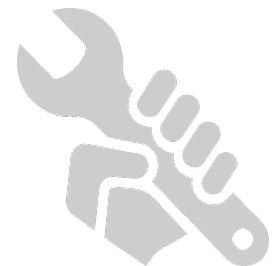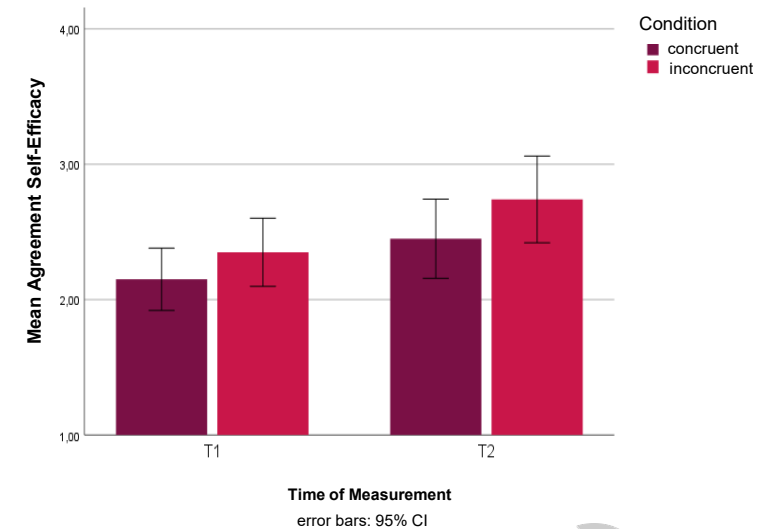
Laboratory: self-efficacy expectation 3/10

Self-efficacy rating [18]

Rating of self-assessed efficacy on a single-item, 100-point rating scale according to Bandura (2006):

"Please rate how confident you are in dealing with privacy or data protection issues in the context of mobile apps. Rate your confidence by choosing a number from 0 to 100 from the scale."

Measurement	n	M	SD	MIN	MAX
Pre-measurement	22	45.23	19.25	13.00	81.00
Post measurement	22	60.14	21.98	11.00	93.00

The data before and after trying out the prototype differ highly significantly:

The self-efficacy of the test subjects could be significantly improved by the interaction with the prototype.

error bars: 95% CI

[18] Bandura (2006); https://iconmonstr.com/wrench-24-png/
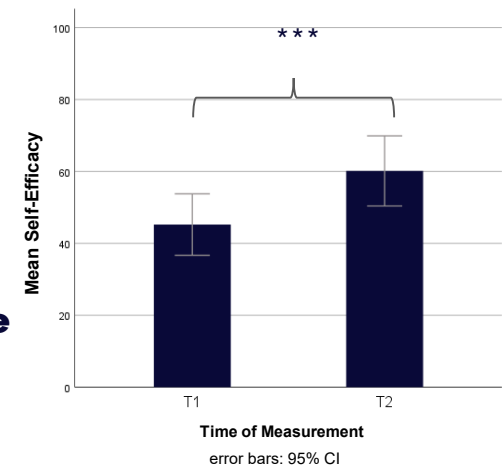
Laboratory: self-efficacy expectation 4/10

Self-efficacy rating [18]

Main effect time but not for experimental condition and not for the interaction time x experimental condition.

The subjects' self-efficacy was significantly improved by their interaction with the prototype. **Which prototype variant (suitable/incompatible) the test subjects tried out again did not play a role.**

error bars: 95% CI

[18] Bandura (2006); https://iconmonstr.com/wrench-24-png/
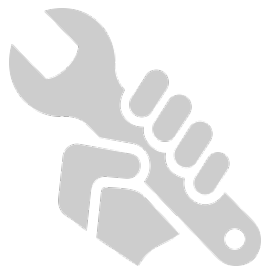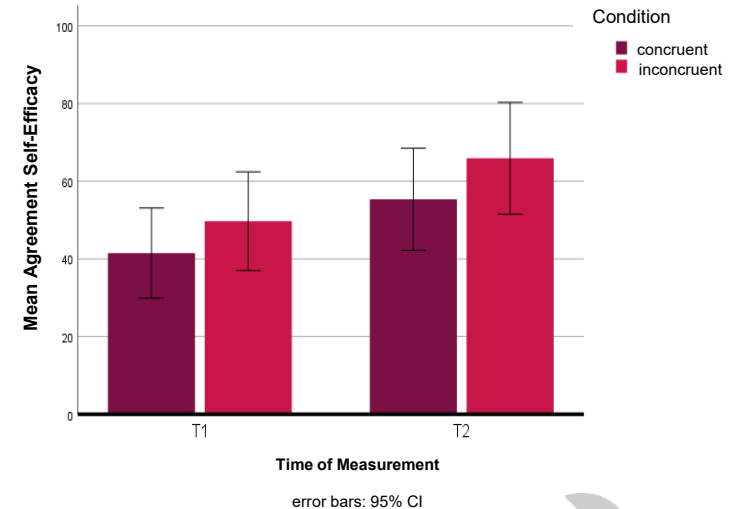
Laboratory: self-efficacy expectation 5/10

Self-Efficacy in Information Security (SEIS) [19]

SEIS (11 items)

Measurement	n	M	SD	MIN	MAX
Pre-measurement	22	4.98	1.11	2.00	6.73
Post-measurement	22	5.46	1.22	1.45	7.00
Follow-Up	22	5.50	1.02	2.64	6.82

Self-efficacy regarding information security increases significantly after interaction with the PANDERAM prototype and remains stable over a short period of time.

error bars: 95% CI

[19] Rhee, H.-S., Kim, C., & Ryu, Y. U. (2009).

50
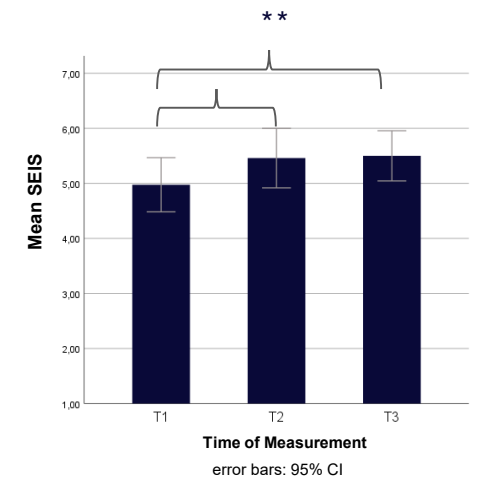
Laboratory: self-efficacy expectation 6/10

Self-Efficacy in Information Security (SEIS) [19]

Defines self-efficacy in information security as the belief in one's own ability to protect information and information systems (smartphone in this case) from unauthorized disclosure, modification, loss, destruction, and lack of availability.

SEIS (11 items)

T2: $Mdn_{concruent}$ = 5.27; $Mnd_{inconcruent}$ = 5.73
T3: $Mdn_{concruent}$ = 5.82; $Mnd_{inconcruent}$ = 5.73

After the interaction with the prototype (T2, T3), the SEIS scores of individuals in the different experimental conditions do not differ significantly, i.e., the **experimental condition does not play a role with regard to self-efficacy for information security.**

[19] Rhee, H.-S., Kim, C., & Ryu, Y. U. (2009).
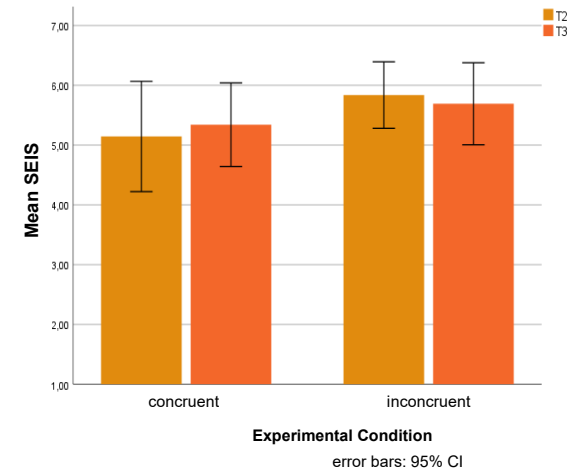
Laboratory: self-efficacy expectation 7/10

Self-Efficacy in Information Security (SEIS) [19]

General Controllability (3 items)

Measurement	n	M	SD	MIN	MAX
Pre-measurement	22	3.80	1.21	1.00	6.00
Post-measurement	22	4.48	1.51	1.00	7.00
Follow-Up	22	4.39	1.44	1.33	7.00

The conviction regarding the controllability of data security rises after the interaction with the PANDERAM prototype and remains (reasonably) stable over a short period of time.

error bars: 95% CI

[19] Rhee, H.-S., Kim, C., & Ryu, Y. U. (2009).
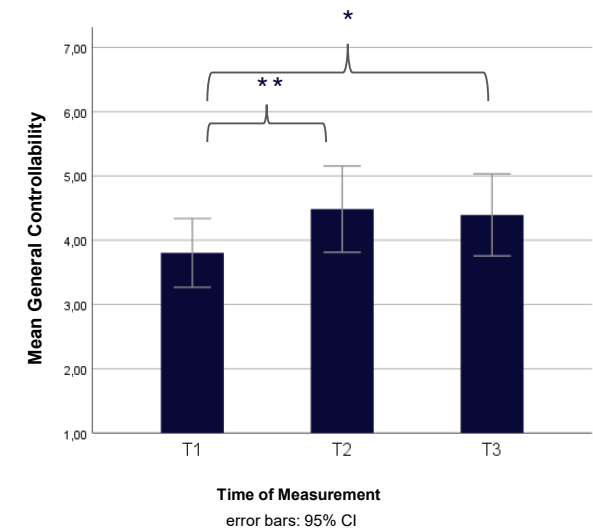
Laboratory: self-efficacy expectation 8/10

Self-Efficacy in Information Security (SEIS) [19]

General Controllability (3 items)

Main effect for time but not for experimental condition or interaction effect.

The conviction regarding the controllability of data security increases after the after interaction with the PANDERAM prototype and remains (reasonably) stable over a short period of time, **but is not dependent on the experimental condition.**

Condition
■ concruent
■ inconcruent

error bars: 95% CI

[19] Rhee, H.-S., Kim, C., & Ryu, Y. U. (2009).
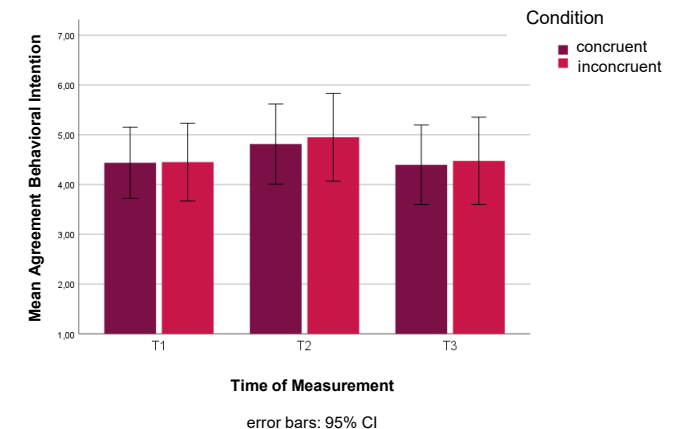
Laboratory: self-efficacy expectation 9/10

Self-Efficacy in Information Security (SEIS) [19]

Behavioral Intention (4 items)

Measurement	n	M	SD	MIN	MAX
Pre-measurement	22	4.44	1.16	1.75	6.25
Post-measurement	22	4.88	1.31	2.50	7.00
Follow-Up	22	4.43	1.39	2.25	7.00

The intention to do something for his information security increases after the after interacting with the PANDERAM prototype, but does not persist over a short period of time.

[19] Rhee, H.-S., Kim, C., & Ryu, Y. U. (2009).
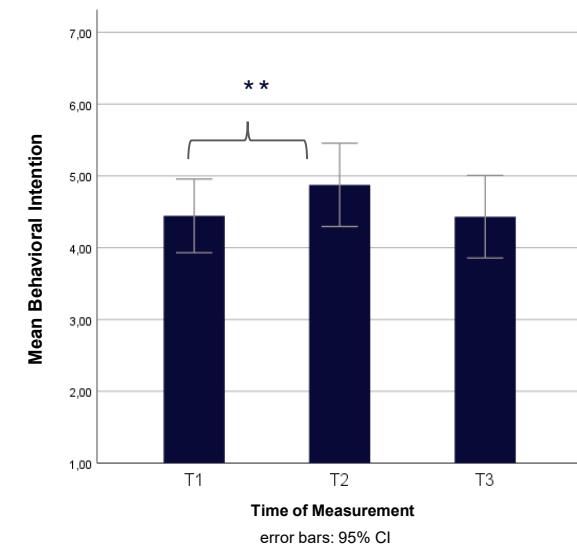
Laboratory: self-efficacy expectation 10/10

Self-Efficacy in Information Security (SEIS) [19]

Behavioral Intention (4 items)

Main effect for time but not for experimental condition or interaction.

The intention to do something for their information security increases significantly after the significantly after interacting with the PANDERAM prototype, but does not persist over a short period of time and **is not dependent on the experimental condition.**

error bars: 95% CI

Condition
■ concruent
■ inconcruent

Time of Measurement

error bars: 95% CI

[19] Rhee, H.-S., Kim, C., & Ryu, Y. U. (2009).
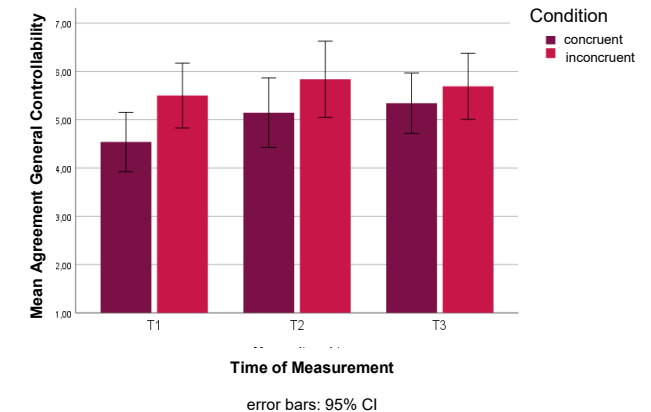
Laboratory: praise for the app (qualitative)

What aspects of the PANDERAM app did you **particularly like**?
Name three aspects that are most important to you.

90 statements were coded into categories with two
parallel levels: 1ˢᵗ level = distinction between "design of
the app" (51% of statements), "information and functions of
the app" (47%), and "other" (2%).

2ⁿᵈ level: more detailed content (see next slide)

"Praise" for PANDERAM app

Design of the app — 51%

Information and function of the app — 47%

Misc. — 2

0% 10% 20% 30% 40% 50% 60%

Relative response frequency

Laboratory: praise for the app (qualitative)

"3. The design is kept very simple and is very clear due to the division into three sections."

Design of the app
- Quick overview
- Color coding risk
- Motivating, animating design
- Simplicity and comprehensibility
- Comprehensive design (intuitive operation, menu navigation, design)

Information and function of the app
- Information on device safety
- Direct (linked) instructions for action
- Information on (app) risks
- Suggestions for app alternatives
- Information on third-party providers (forwarding, processing)

Misc.
- Numerical coding of the risk
- Subject of the app

"that in addition to the content information, there was also color coding that made changes quickly visible."

"Links with settings and not first own switching into other interface".

"... I was also able to see which apps were secure and which ones might also have too many permissions and ..."

Relative response frequency

0%   5%   10%   15%

Laboratory: Criticism for the app (qualitative)

What **opportunities for improvement** do you see for the PANDERAM app?
Name three aspects that are most important to you.

50 statements were coded into categories with two
parallel levels: 1st level = distinction between
"**comprehensibility**" (36% of statements),
"**presentation and design**" (32%). and
"**enhancements"** (22%).
6% of the answers were for "other" and 4% for "no"
suggestions for improvement.

2nd level: more detailed content (see next slide)

Room for improvement PANDERM app

Comprehensibility	36%
Presentation and design	32%
Enhancements	22%
Misc.	6%
No criticism	4%

0%   10%   20%   30%   40%

Relative response frequency

Laboratory: criticism for the app (qualitative)

"...I would advise replacing the numbers with words like "critical/high risk"."

Comprehensibility
- Number coding Risk (scale, value range, sorting)
- Misleading design (colors, menu navigation)
- Terms and explanations unclear, incomprehensible
- Non-transparent risk assessment for third-party providers/apps

3) Especially in "Device Security" I don't find the subdivision into the categories useful, because you actually don't know where to find what: why is the lock screen ei User Settings and not in Device Settings? I would find it easier here if all the items were simply listed one after the other. Otherwise, it needs to be clearer what can be found under each menu item. "

Presentation and design
- Links for all options for action
- Reduce/restructure information content
- General design (colors, menu navigation, modernity, motivation)

"For security updates and fingerprint, no action buttons to change these - you have to look in the device settings first."

"Optics/layout/color scheme could be made a bit more appealing or "modern" in parts (however, no criticism of color scheme regarding risk levels, as these are good + generally understandable and should be retained)."

Enhancements
- More information and functions for advanced users
- Further information on this topic

"1. there is too little attention to advanced techniques to increase safety. The app is aimed at beginners, so that advanced users have little fun. Password managers, VPN browsers, deleting cookies, etc. should be mentioned here. "

Misc.

No criticism

0%　5%　10%　15%　20%

Relative response frequency

Laboratory: School grade

What grade would you give the PANDERAM app overall?

On a scale of (1 = *very good*, 2 = *good,* 3 = *satisfactory,* 4 = *sufficient*, 5 = *poor,* and 6 = *unsatisfactory*), the *n* = 21 participants in the laboratory experiment gave an average grade of "***good.***"
(*Mdn* = 2.00; *MW* = 1.71; Min = 1.00; Max = 3.00).

The experimental condition or the behavioral stage did not any difference for the grading.

n = 21
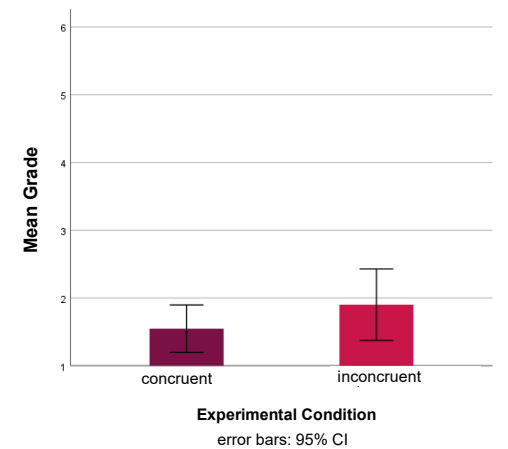
Laboratory: interim conclusion

The usability rating of the PANDERAM app was **good** (as in laboratory test 1 and despite an increase in content).

- Whether the participants were presented with the appropriate or inappropriate variant for the behavioral stage is irrelevant.

The **user experience is** rated as **excellent.** Only in terms of attractiveness and originality do the participants see a slight need for improvement.

- Tailoring has an effect here for the evaluation of the aspect "novelty", i.e. the **suitable app variant arouses a higher interest** than the unsuitable one.

Overall, significantly **more aspects of the app** were **praised** than criticized, especially:

- the **overview and clarity that** the app provides through its simple design.
- the **direct linking of action options to** the settings in the smartphone.
- the **color scheme** (traffic light colors) of the **risk**

The participants saw the following as areas in need of improvement:

- the **general design of** the app (modernity & motivation),
- the **numerical coding of the risk,** and
- the **lack of focus on experts**

Overall, the **PANDERAM app** was **rated positively** and with the school grade **"good"**.

Laboratory: interim conclusion

(Only) for individuals of **early behavioral** stages, the **appropriate variant of** the app **increases concern** and can thus contribute to the **creation of problem awareness**, which is a prerequisite for protective behavior.

For effort expectancy and sense of control, tailoring the app had **no effect,** but interaction with the app itself or the **temporal difference** due to measurement repetition did.

On all scales used, there is a **clear and stable short-term increase in self-efficacy** through the PANDERAM app, regardless of the variant presented.
It thus contributes to becoming **more self-confident in dealing with privacy or data protection problems** when using smartphones and apps.

In addition, the app **reinforces the** belief in the **controllability of threats** and the **behavioral intention to take measures** to increase security.

Follow-up: questions & hypothesis

Q3: Do changes in intention to increase smartphone security occur as a result of using the PANDERAM app?"

Tends to have later behavioral stages at follow-up than at recruitment.

Behavioral stages **summarized:**

- remained the same: 10 (of which 8 Postaction)

- back (late → early): 2

- on (early → late): 10 (of which 7 to Postaction)
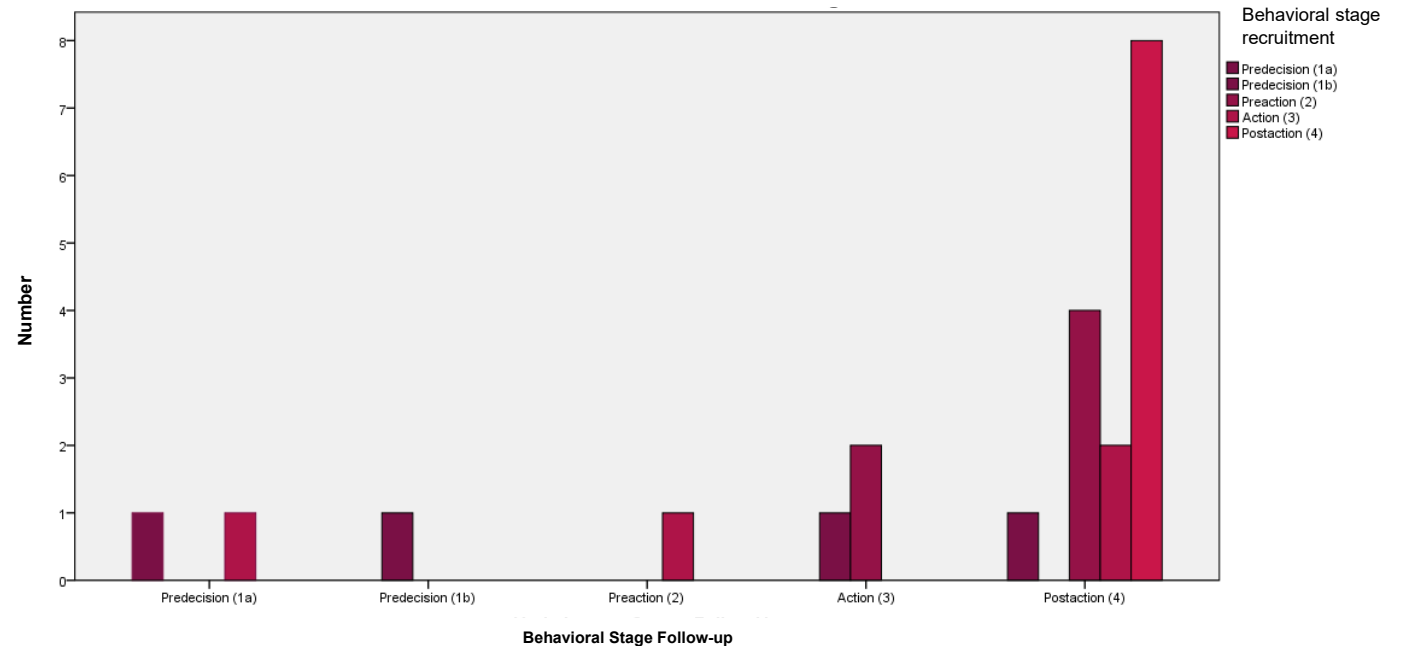
Follow-up: questions & hypothesis

Q3: Does the use of the PANDERAM app result in any changes in the intention to increase smartphone security?"

General behavioral (intention) change: (n = 22).

50% (11) "*No, nothing has changed*",

32% (7) "*Yes there has been some change*" and 18% (4) "*I'm not sure*",

Of which 18% (2) were "*app related*", 45% (5) were "*smartphone related*" and

37% (4) "*App and smartphone related*".

→ **9 of the 11 participants who answered** "*Yes*" or "*Not sure.*" answered "Yes" or "Not sure" have indicated behavioral (intention) changes to **actions**, which were **also to be** were to be processed.

General change
in behavior
(intention)

■ Yes, something has
changed.
■ No, nothing has
changed.
■ I'm not sure.

32%
50%
18%

Number

Yes   No

Named changes have been part of the lab test
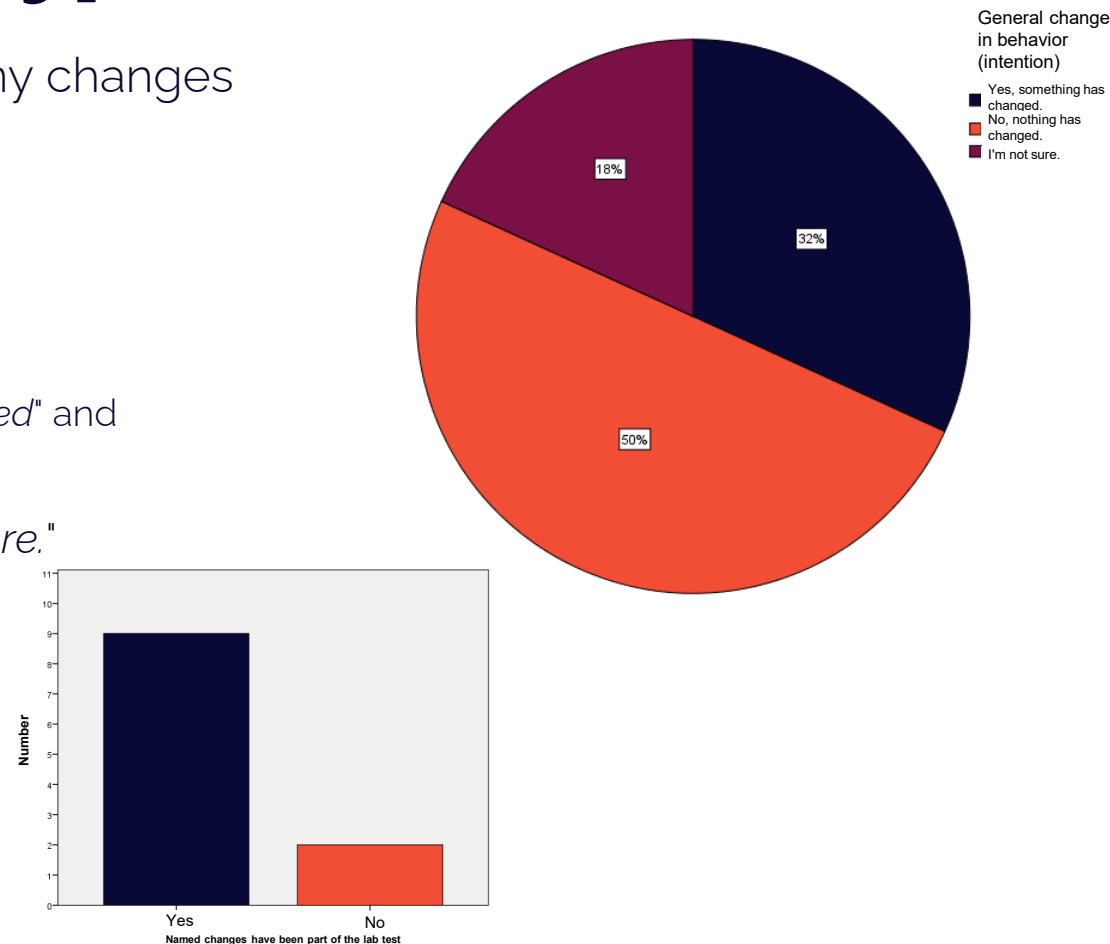
Follow-up: changes (qualitative)

10 days ago, you completed the test in our laboratory. In the process, you became intensively acquainted with the PANDERAM app and its contents.

Did you notice any changes in your smartphone-related **privacy behavior/attitude toward smartphone-related privacy/attention to smartphone-related privacy** after the trial?

Fifty-five percent (13) of subjects provided information on changes in at least one of these questions.

The descriptions contained **44 individual aspects**.

Behavioral changes (41%) were described most frequently, followed by changes in attention (34%) and changes in attitude (+knowledge, sense of control, and sense of competence).

Changes after laboratory test

Behavioral change	41%
Change in attitude (incl. increase in knowledge, perception of control/competence)	23%
Change in attention	36%

0,00% 10,00% 20,00% 30,00% 40,00% 50,00%

Relative response frequency

Follow-up: praise for the app (qualitative)

Which aspect of the PANDERAM app do you **remember most positively**? Name the aspect that is most important to you.

Thirty-four statements were coded into categories with two parallel levels: 1st level = distinction between "design of the app" (50% of statements), "information and functions of the app" (44%), and "subject of the app" (6%).

2nd level: more detailed content (see next slide)

Positive aspects (T3)

Design of the app	50%
Information and functions of the app	44%
Subject of the app	6%

0%   20%   40%   60%

Relative frequency of response

67

Follow-up: praise for the app (qualitative)

"...get an overview of the state of security."

"the intuitive operation of the interface

"small explanations / information why certain aspects are considered risky "

"Direct redirection to system settings to make changes easily without much clicking."

"The idea behind the app to inform, educate and help people improve their data security."

Design of the app
- Clear arrangement and overview
- Intuitive operation and simplicity
- Design of the risk assessment (colors)

Information and functions of the app
- Information and explanations (general)
- Direct forwarding to settings
- Risk assessment
- App alternatives
- Information on third-party providers

Subject of the app

0% 5% 10% 15% 20% 25%

Relative response frequency

Follow-up: criticism for the app (qualitative)

Which aspect of the PANDERAM app do you remember as **needing particular improvement**? Name the aspect that is most important to you.

There were **28 statements** in 4 categories:

1.) Misleading risk information (21% of statements),

2.) Design of information and functions (50%),

3.) Extended information (7%) and

4.) no aspects (21%)

Improvement possibilities PANDERAM app (T3)

Category	Value
Misleading risk information (scale, colors, comprehensibility)	21%
Design of information and functions of the app (design, motivation, clarity, menu navigation, accessibility)	50%
Extended information for experts and further information	7%
Nothing	21%

0% 10% 20% 30% 40% 50% 60%

Relative response frequency

69
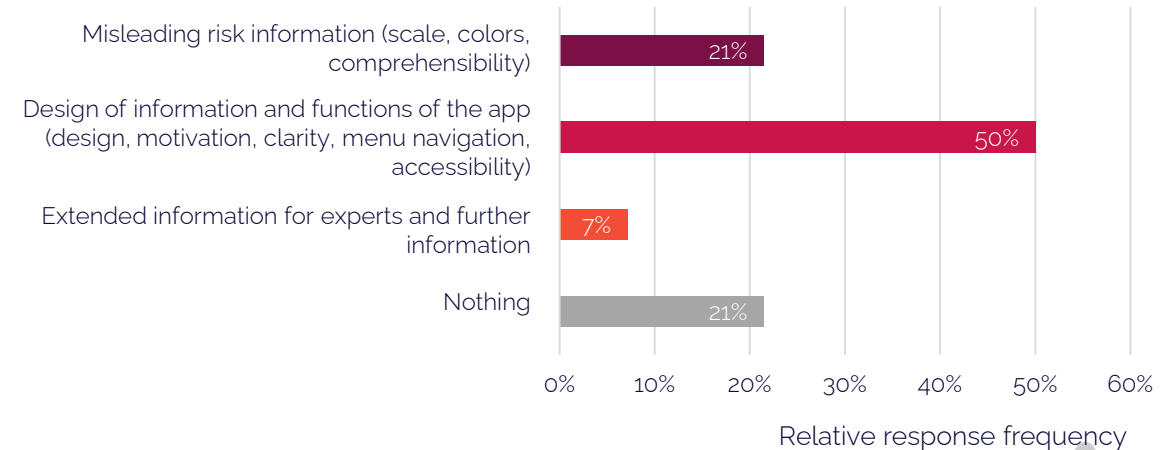
Follow-up: relevance of the app (quantitative)

To what extent do you consider the PANDERAM app and its content to be...

)

...personally relevant?

...socially relevant?

M_{pers} = 4.36 (SD = .79) vs. M_{soc} = 3.73 (SD = 1.16)
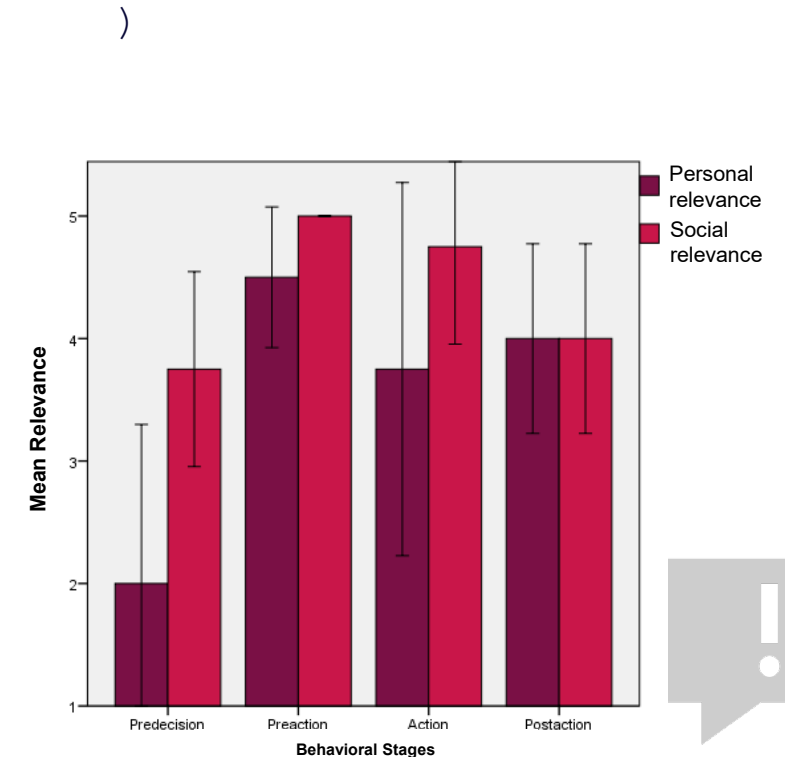
Significant differences between behavioral stages.

error bars: 95% CI

Follow-up: personal relevance of the app

To what extent do you consider the PANDERAM app and its content to be personally relevant? Briefly justify your statements on personal relevance in your own words.

There were **27 statements** on:

1.) high personal relevance (78%),

2.) low personal relevance (19%), and

3.) Other aspects (4%) and

coded with two parallel levels.

2nd level: More detailed content (see next slide).

Statements on personal relevance of the PANDERAM app (T3)

Personal relevance high (>2) — 78%

Personal relevance low — 19%

Misc. — 4%

0%   50%   100%

Relative response frequency

Follow-up: personal relevance of the app

Personal relevance high (>2)
- Personal interest in data protection
- Attention & sensitization (for improvements in data protection of own data and in general)
- Real security of your own data
- Increase knowledge & skills in dealing with data protection (education, information)
- Motivate & activate for action
- Facilitation through Panderam app (get an overview, save time)

Personal relevance low
- Not dealing with data protection issues (ignoring relevance, having no influence)
- No benefit or gain from the Panderam app

Misc.

"My data is important and can be misused in many ways"

" I myself would install the app, even though I have already paid attention to many apps before what permissions they have. I think it would either confirm me then dari or show me places where I can still improve something "

"As I have not been much concerned with the security of my smartphone regarding my privacy, the simple explanations were very enlightening and educational for me/my personal knowledge level

"I already know some things about data security, so I have little need for the app."

Relative response frequency

0%   5%   10%   15%   20%
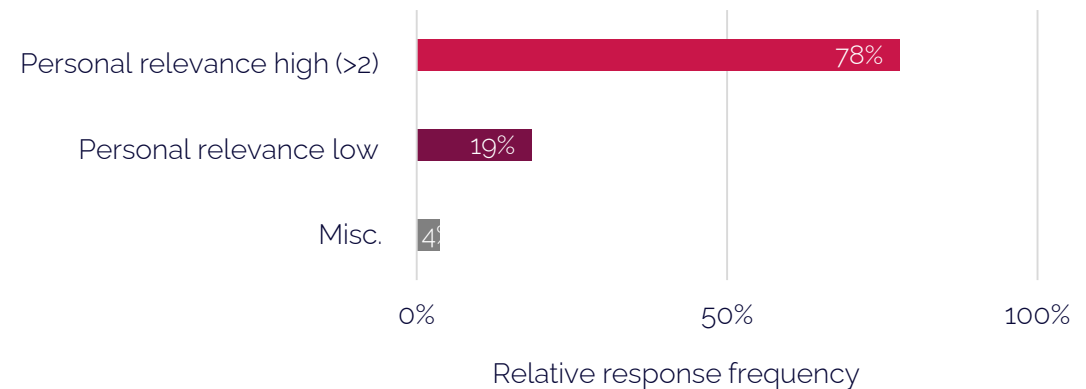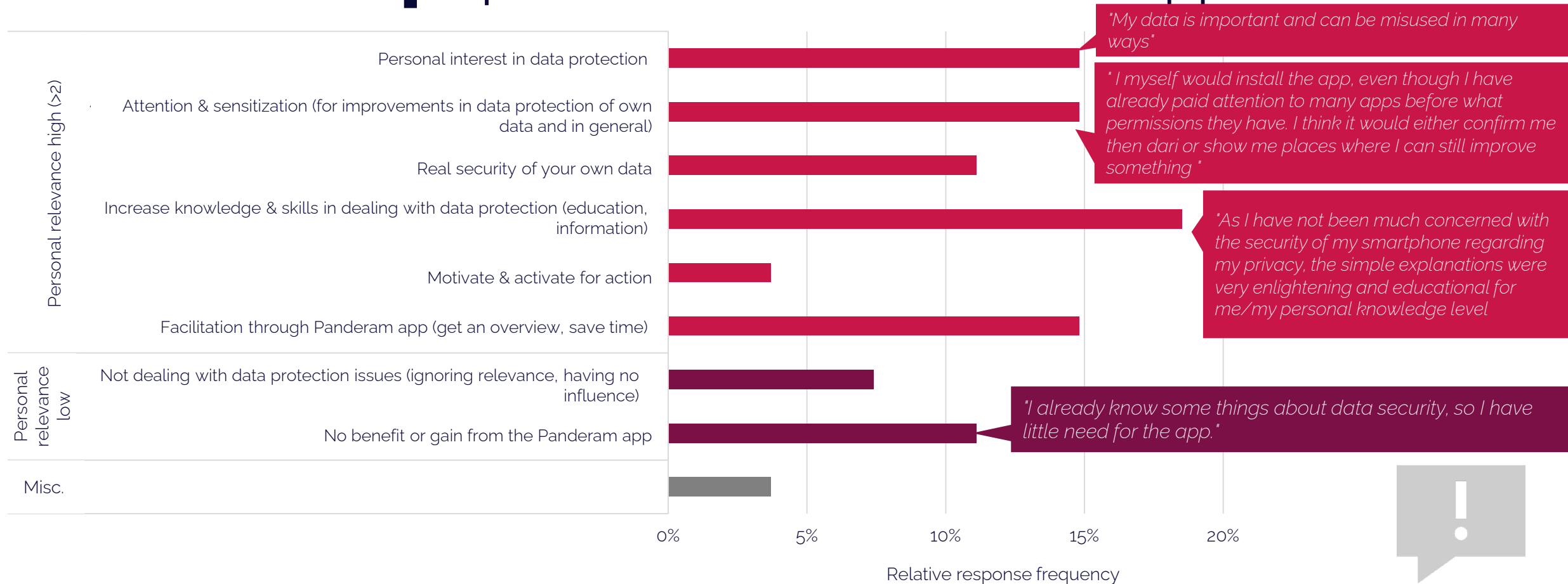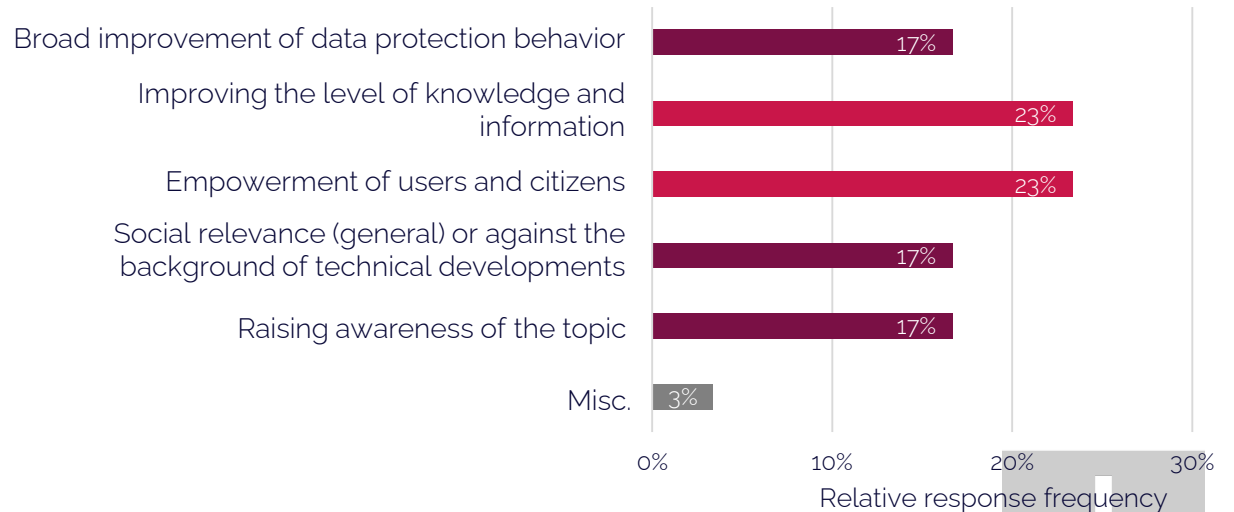
Follow-up: social relevance of the app (qualitative)

To what extent do you consider the PANDERAM app and its content to be socially relevant? Briefly justify your statements on social relevance in your own words.

There were **30 statements** assigned to 6 categories most frequently:

1) Empowerment of users and citizens (23%),

2) Improvement of knowledge and informedness (23%).

Statements on the social relevance of the PANDERAM app (T3)

Category	Relative response frequency
Broad improvement of data protection behavior	17%
Improving the level of knowledge and information	23%
Empowerment of users and citizens	23%
Social relevance (general) or against the background of technical developments	17%
Raising awareness of the topic	17%
Misc.	3%

Relative response frequency

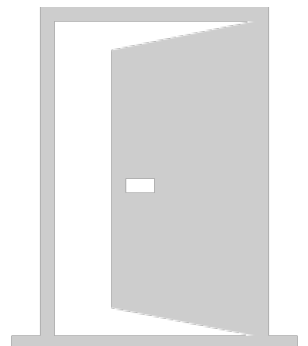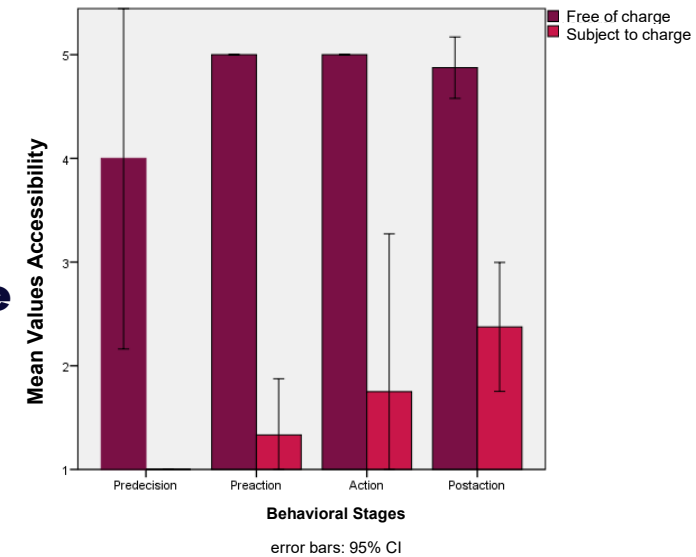# **Follow-up:** accessibility of the app (quantitative)

*"The PANDERAM app should be **public and free for all to use**." vs.*

*"The PANDERAM app should be available for use by **interested parties for a fee**."*

Range: 3 - 5 (consistently medium to high agreement) vs.
Range: 1 - 3 (consistently low to medium agreement)

M_{free} = 4.77 (SD = .612) vs. M_{fee} = 1.73 (SD = .827).

Behavioral stages differed in their agreement for the paid option.

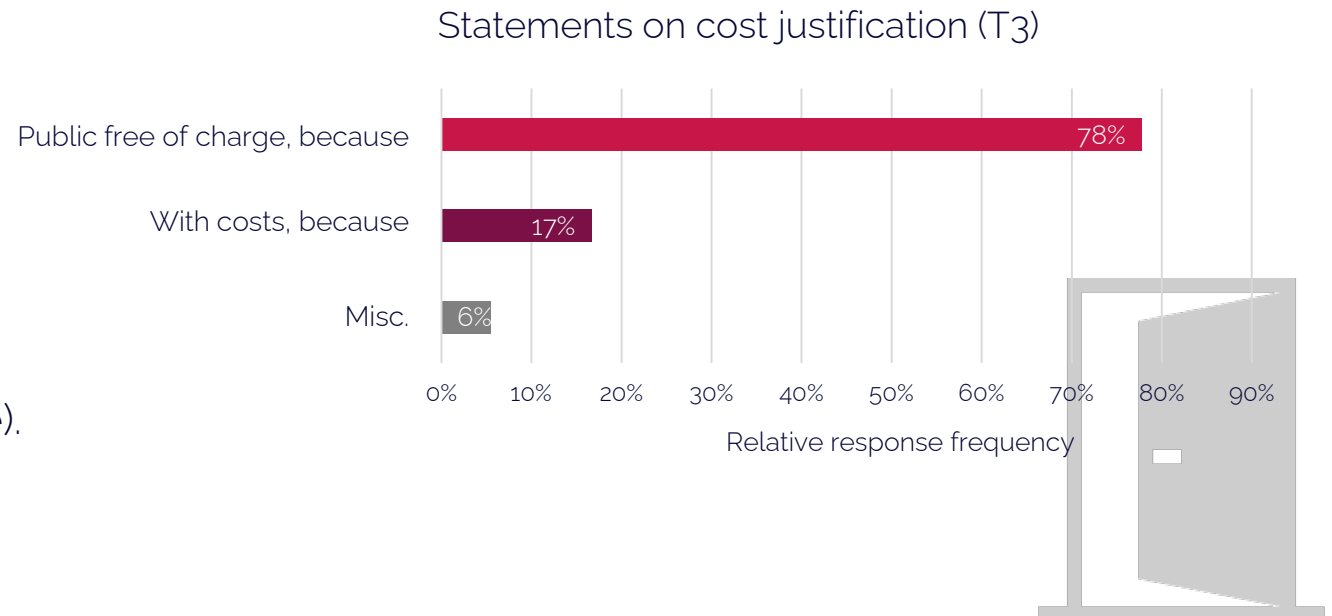# **Follow-up:** accessibility of the app (qualitative)

"The PANDERAM app should be public and free for all to use." vs. *"The PANDERAM app should be available for use by interested parties for a fee."*
Briefly justify your statements in your own words.

There were **36 statements** categorized as:
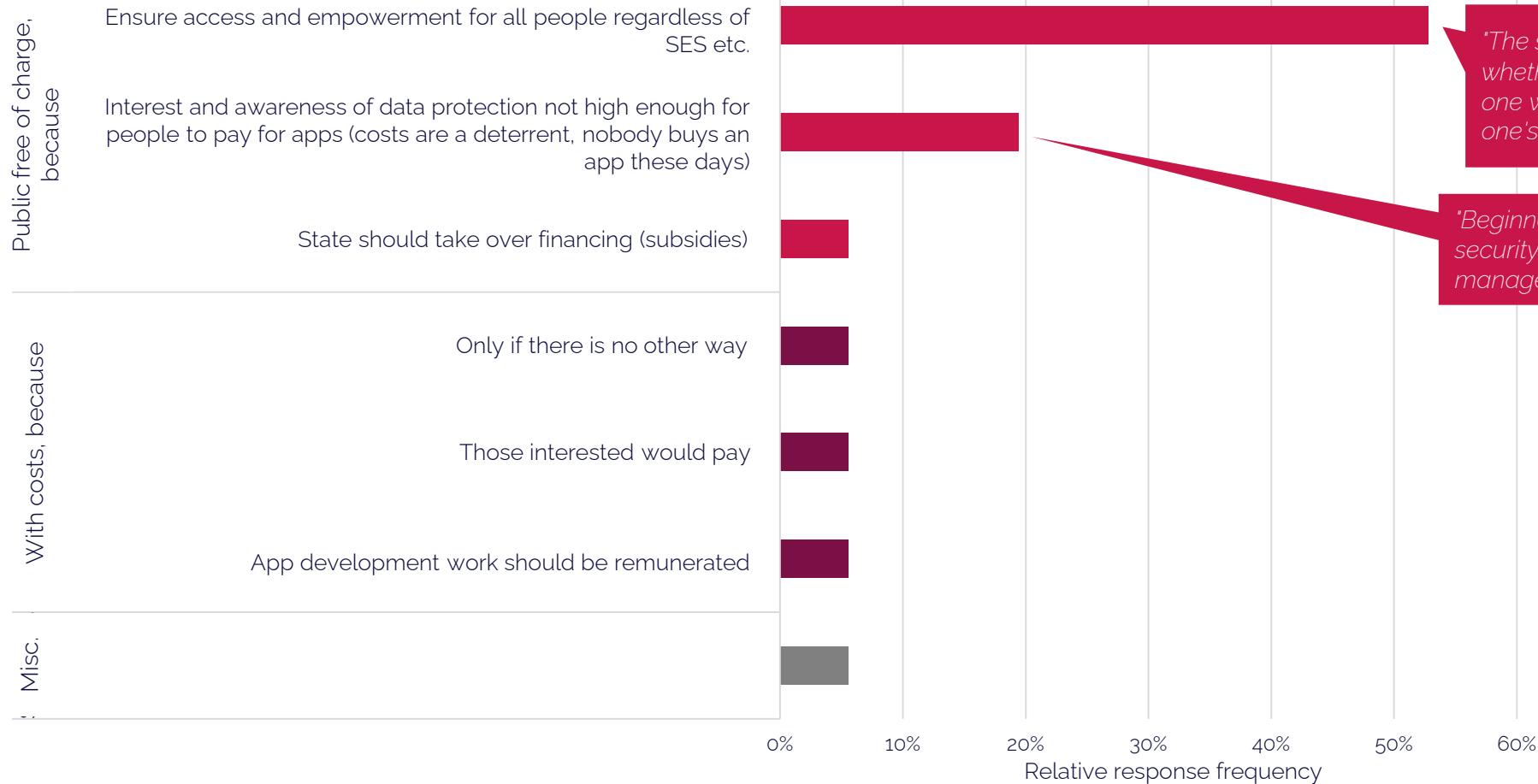
1.) Publicly accessible (78%),

2.) Chargeable (17%), and

3.) Other aspects (6%) and

coded with two parallel levels.

2nd level: More detailed content (see next slide).

Statements on cost justification (T3)

Public free of charge, because — 78%

With costs, because — 17%

Misc. — 6%

0% 10% 20% 30% 40% 50% 60% 70% 80% 90%

Relative response frequency

Follow-up: accessibility of the app (qualitative)

Public free of charge, because

- Ensure access and empowerment for all people regardless of SES etc.
- Interest and awareness of data protection not high enough for people to pay for apps (costs are a deterrent, nobody buys an app these days)
- State should take over financing (subsidies)

With costs, because

- Only if there is no other way
- Those interested would pay
- App development work should be remunerated

Misc.

"The security of one's own data should not depend on whether one can afford the app, but solely on whether one wants to protect one's data or not. The security of one's data should therefore be free. "

"Beginners tend to have little interest in the topic of data security, and a paywall (as is often the case with password managers) would be more of a deterrent. "

Relative response frequency

0%   10%   20%   30%   40%   50%   60%
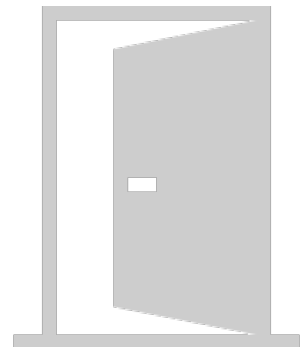
Follow-up: interim conclusion

Approximately 10 days after the laboratory test, **more than 50% of** the participants stated that they had noticed **differences in privacy protection behavior and behavioral intention.** This is also reflected in the reported changes in topic-specific **attention or attitude,** as well as the assignment to **behavioral stages**, which shifted in **favor of the later behavioral stages.** The behavioral changes described are similar to the tasks that were completed by the participants in the **laboratory experiment.**

After the laboratory test, therefore, **at least half of the participants experienced a change in their behavior, intention, attention or attitude toward smartphone-related privacy protection.** A causal relationship with the use of the PANDERAM app cannot be statistically proven, but seems likely.

The PANDERAM app was also deemed to have **high personal and social relevance,** and **public, free accessibility was** demanded.

Summary of laboratory test 2

- People in the **early behavioral stages consider** themselves **less concerned about** smartphone safety than people in the late behavioral stages.

- A **variant** of the app **that matches** the behavioral stage, **increased** this **concern** of the early behavioral stages and can thus create a basis for adequate problem awareness and behavioral change.

- With regard to **perceived control** and **expected effort** for (increasing) smartphone security, **no differences were** found between people of different behavioral stages. The behavioral stage of the app also had **no effect** here.

- The participants rated the app as **highly usable** and attested to a very good user experience.
  a **very good user experience.**

- The strengths of the app lie in the overview that the app provides,
  the directly linked options for action and the color coding of the risk.

- There is a slight need for improvement with regard to the modern and motivating design,
  the numerical coding of the risk, and the focus on experts.

Summary of laboratory test 2

- The PANDERAM app makes a clear and short-term stable contribution to **increasing self-efficacy in dealing with privacy and data protection issues.**

- It strengthens the belief in the controllability of security threats and the intention to take measures to increase smartphone security.

- The **majority of users** reported **changes in attention, attitude and behavior** after interacting with the PANDERAM app.

- The app is also considered to have a **high level of personal and social relevance.**
  From the participants' point of view, it should be available to the public and free of charge.
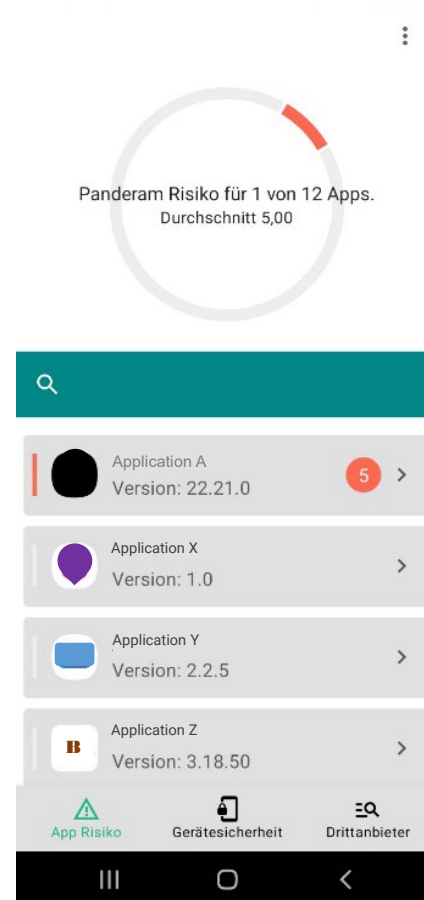
Summary

The PANDERAM app...

...can bring about change for half of the users.

...is relevant.

...makes users more self-effective.

...is designed in a user-friendly way.

...should be expanded for experts.

...and the implemented behavioral stage tailoring had only limited effects.

Panderam Risiko für 1 von 12 Apps.
Durchschnitt 5,00

Application A
Version: 22.21.0
5

Application X
Version: 1.0

Application Y
Version: 2.2.5

Application Z
Version: 3.18.50

App Risiko   Gerätesicherheit   Drittanbieter

Outlook

The second laboratory test represents the conclusion of the sub-project User-Centered Design of the Chair of General Psychology and Human Factors at Chemnitz University of Technology.

The collected results will be included in the final report of the PANDERAM project.

Thank you for your attention!

Susen Döbelt

Wilhelm-Raabe-Str. 43
09120 Chemnitz

Phone: 0371 531 33615
E-mail: susen.doebelt@psychologie.tu-chemnitz.de