Out-of-Device Privacy Unveiled: Designing and Validating the Out-of-Device Privacy Scale (ODPS)

Habiba Farzand
h.farzand.1@research.gla.ac.uk
University of Glasgow
United Kingdom

Karola Marky
karola.marky@rub.de
Ruhr University Bochum
Germany

Mohamed Khamis
mohamed.khamis@glasgow.ac.uk
University of Glasgow
United Kingdom

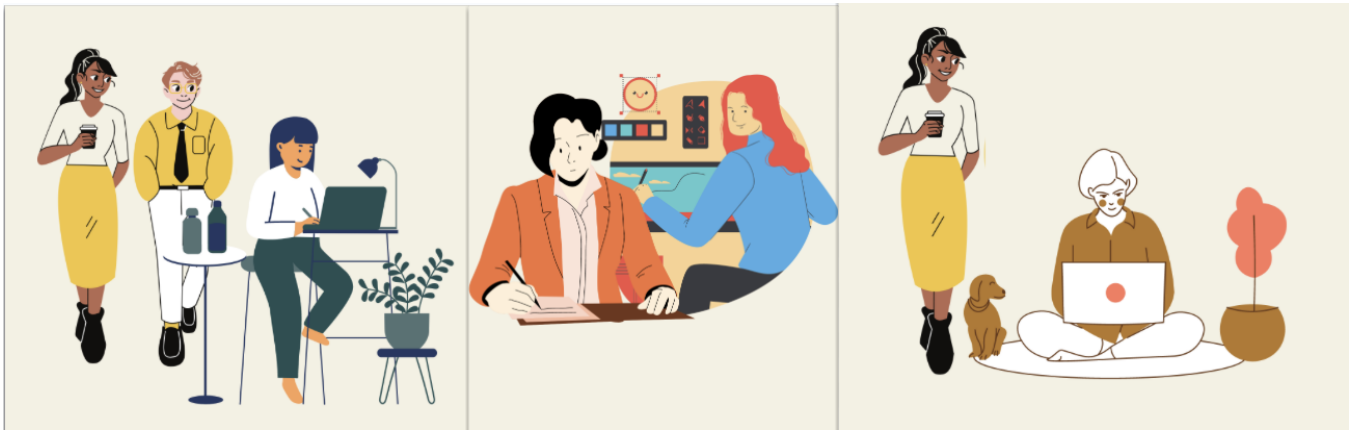

Figure 1: The figure shows daily life scenarios when out-of-device privacy threats in the physical world, such as shoulder surfing, take advantage of the user's physical surroundings to invade data privacy without the user realizing it. (The image was created using Canva under free license [13].)

ABSTRACT

This paper proposes an Out-of-Device Privacy Scale (ODPS) - a reliable, validated psychometric privacy scale that measures users' importance of out-of-device privacy. In contrast to existing scales, ODPS is designed to capture the importance individuals attribute to protecting personal information from out-of-device threats in the physical world, which is essential when designing privacy protection mechanisms. We iteratively developed and refined ODPS in three high-level steps: item development, scale development, and scale validation, with a total of N=1378 participants. Our methodology included ensuring content validity by following various approaches to generate items. We collected insights from experts and target audiences to understand response variability. Next, we explored the underlying factor structure using multiple methods and performed dimensionality, reliability, and validity tests to finalise the scale. We discuss how ODPS can support future work predicting user behaviours and designing protection methods to mitigate privacy risks.

CHI '24, May 11–16, 2024, Honolulu, HI, USA

ACM ISBN 979-8-4007-0330-0/24/05
https://doi.org/10.1145/3613904.3642623

CCS CONCEPTS

• **Security and privacy** → **Economics of security and privacy**; **Privacy protections**; **Social aspects of security and privacy**;

KEYWORDS

out-of-device privacy, out-of-device threats, privacy in the physical world, scale development, privacy

ACM Reference Format:
Habiba Farzand, Karola Marky, and Mohamed Khamis. 2024. Out-of-Device Privacy Unveiled: Designing and Validating the Out-of-Device Privacy Scale (ODPS). In *Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI '24), May 11–16, 2024, Honolulu, HI, USA.* ACM, New York, NY, USA, 15 pages. https://doi.org/10.1145/3613904.3642623

1 INTRODUCTION

"There are things known, and there are things unknown, and in between are the doors of perception."

Aldous Huxley, 1954

Threats to information privacy are not only restricted to device use, such as GUI confusion attacks [8] since the technology surrounding us continuously collects sensitive information and can be used maliciously. Everyday scenarios, such as withdrawing cash at an ATM and being recorded by CCTV [15, 61] or travelling on a bus and being shoulder surfed [22, 25, 26] - all these scenarios make the user's data susceptible to attacks in the physical world. As we transition into a society increasingly reliant on technology, privacy concerns are becoming more pervasive.

While some attacks on user privacy require the advanced, sophisticated expertise of the attacker, the increased usage of (mobile) devices and tools has enabled such attacks with only little expertise. For example, shoulder surfing can be done through direct observation [22] or recording videos [66]. Moreover, user privacy can be violated simply by using technology gadgets, such as a thermal camera [3] that can infer sensitive input entered on keyboards [1, 2]. The combination of little required expertise for attack execution and increased availability of resources amplifies the vulnerability to privacy attacks. Due to this, anyone could invade anyone's privacy, putting everyone's privacy at risk. Privacy researchers have proposed numerous mechanisms to mitigate the out-of-device privacy threats in the physical world. Such mechanisms include visual filters, generating fake text, icon overlaying and vibratory alerts [39, 56, 68]. However, the one-size-fits-all approach cannot be applied. For example, shoulder surfing - a privacy threat that exists out of the device is perceived as concerning by some people, whereas some people do not consider shoulder surfing a risk [22, 31]. In contrast, others switch off or cover the device with their hand when suspecting a shoulder surfing attack [22, 25, 32]. The differences in reaction towards out-of-device privacy invasions reflect the differences in users' out-of-device privacy profiles. Therefore, precisely mapping mechanisms to user profiles is challenging without knowing users' privacy profiles. Safeguarding privacy from out-of-device threats requires investigating how much importance users prescribe to such threats defined by how users use tech and respond to privacy violations. We propose utilising the notion of *"out-of-device privacy"* to capture this.

While literature lists several privacy-related measures like IUIPC [44], they are limited to specific scopes; for example, the IUIPC precisely measures online information privacy concerns. To date, there is no standard scale to measure users' associated importance towards threats in the physical world. To explore users' importance towards protecting personal information from threats in the physical world, we first propose a definition for *"out-of-device privacy"* and second present an 18-item psychometric scale instrument to measure the importance a person attributes to protecting personal information from out-of-device threats in the physical world. Our scale development process involved three steps: (1) item development, (2) scale development, and (3) scale validation. In the item development phase, we aimed for content validity by following deductive and inductive approaches to collect an initial item pool from literature and experts (N=13). Next, we pre-tested the items with experts and the target audience (N=48), which assisted in refining the wording and provided initial insights into the variability of responses. Finally, we deployed the survey online (N=382) in the scale development phase and used the data to extract underlying factors. Lastly, we performed dimensionality, reliability, and validity tests on a dataset of N=935 participants in the scale validation phase. The scale was iteratively developed and refined throughout all stages in multiple studies involving N=1378 participants. Finally, we confirmed the scale structure and presented the 18-item Out of Device Privacy Scale (ODPS).

Our scale establishes a foundation for assessing out-of-device privacy, facilitating systematic analysis and comparison. The scale provides a lightweight method for security and privacy researchers and technology developers to evaluate and predict users' behaviour to protect the users' data from out-of-device privacy threats in the physical world.

2 BACKGROUND

This section overviews existing privacy scales and discusses the research gap.

2.1 Measures of Privacy

Information *privacy* refers to one's desire to control data related to access, use, and sharing [6]. Privacy has been thoroughly investigated in the psychology literature, and numerous attempts have been made to define and measure it. Questionnaires such as IUIPC [44], Privacy Attitude Questionnaire (PAQ) [16], Westin's Privacy Segmentation Index [36], Concern for Information Privacy (CIFP) [62], Global Information Privacy Concern (GIPC) [44], Online Privacy Concern [10] are among the popular privacy scales in the literature.

Internet Users' Information Privacy Concern (IUIPC) [44] measures the information concern of internet users. While this is a reliable and valid instrument, it is limited to internet users only. However, non-internet applications also require consideration of user privacy, such as photos in the phone gallery app. Alan Westin presented Westin's Privacy Segmentation Index to measure privacy perspectives over time; however, the scale only focuses on the organisation's collection and handling of information. Similarly, the Global Information Privacy Concern Scale covers privacy issues related to online companies but lacks evidence of how the statements comprising the scale were selected. The Concern for Information Privacy (CFIP) [62] scale considers consumer online privacy but it lacks the definition of concern. However, it is a well-validated instrument that can only be used in consumer online privacy contexts. Further, the Privacy Attitude Questionnaire (PAQ) [16] considers the privacy concept as a whole and is unsuitable for measuring a specific attribute of privacy. Most importantly, the scale questionnaires mentioned above only consider online information privacy, not privacy in the physical world.

To sum up, the scales detailed above either focus on a specific attribute or capture the concept of privacy as a whole. Moreover, there is a lack of definitional clarity regarding the objectives of some scales. Finally, these scales only focus on internet use and do not consider privacy in the physical world. Regarding privacy invasions in the physical world, physical world elements, such as awareness, influence the user preferences for protection [23–25]. There is a need, therefore, for a validated psychometric instrument to measure and capture people's out-of-device privacy.

2.2 Out-of-Device Privacy in the Literature

Oates et al. [50] conducted a study to explore differences between the privacy mental models of experts and laypersons by asking an open-ended question to express what privacy means to them. Most participants expressed opinions about privacy in the physical world. This indicates that while protecting user's privacy online is essential, protecting it in the physical world is equally important. Further, in a related research study by Gerber et al. [28], they found that most participants are unaware of the effects of privacy violations and that the users perceive most privacy protections

as too fatiguing and complicated. This might be due to the differences in individual perceptions and needs. This highlights the need for user's personalized privacy protection measures. To offer customised privacy protection to users, we must first understand their expectations and preferences for privacy.

Further, assessments of individual attacks have shown that users are impacted negatively due to privacy violations in the physical world. For example, Eiband et al. [22] reported that shoulder surfing gave rise to awkward situations among users. Similarly, further studies have shown that shoulder surfing causes awkwardness and discomfort and has resulted in interaction time wastage with the device and provided evidence that users are likely to adopt a privacy safeguarding mechanism [25, 27]. Following the same line of research, Farzand et al. [24] proposed a typology of perceived sensitive content in response to the users' accounts of shoulder surfing. However, it only provides a list of content types that require protection in different locations but does not consider the user's privacy profile. Cross-culture examinations of privacy concerns have revealed that privacy violations such as observing someone's screen without permission have severe consequences in the Eastern world compared to the Western world [57]. Muslukhov et al. [47] investigated users' concerns about unauthorized access and reported that participants were highly concerned about insiders (e.g. friends) having unauthorized access to their devices. This shows that privacy violations and concerns are found in public places and private environments, such as one's home. The work mentioned above illustrates the significance of addressing privacy threats in the physical world and designing countermeasures that suit individual needs and requirements.

Apart from designing countermeasures, one way of mitigating privacy risks is through user awareness. Users' awareness and knowledge of privacy threats assist them in better protecting and handling their data. With the spread of awareness information about protecting data from online threats, users have become more conscious of how they handle the information. For example, in a study by Jiang et al. [37], only 14.1% were unaware of malware-based threats. On the contrary, out-of-device privacy threats, such as shoulder surfing, often go unnoticed [29], and users remain unaware of the privacy invasion. Likewise, while people are aware of emerging technologies such as thermal cameras, they don't always envisage these technologies in the context of privacy bypasses [5]. To improve the awareness of privacy threats in the physical world, there is a need to systematically capture users' out-of-device privacy so that adequate awareness plans and evaluations can be conducted.

A reliable and standardized method is needed to capture users' out-of-device privacy information. Such an instrument would measure the out-of-device privacy perception of users that will assist in the design of personalized privacy settings, which would offer an appreciable user experience while maintaining user privacy from privacy threats in the physical world. The scale would also benefit developers in developing privacy-aware technologies, attracting more users to adopt technology as they adopt technologies devoted to protecting their privacy [20].

3 STAGE 1: ITEM DEVELOPMENT

This section describes our iterative approach to developing and refining the items to be included in the out-of-device privacy scale.

3.1 Identification of Construct

Our goal was to develop a scale that assesses out-of-device privacy. Towards developing the scale, our first step was to identify the construct and develop a precise definition using simplistic terms. For creating a construct definition, it is essential to consider that it reflects a measurable concept and is sufficiently distinct from other definitions of related constructs. For this, one researcher defined out-of-device privacy. Next, two researchers discussed the definition and iteratively refined it in multiple rounds. This refinement process resulted in the basic definition of out-of-device privacy that we define as follows:

"the importance a person attributes to protecting personal information from out-of-device threats in the physical world"

3.2 Initial Item Pool Generation

After establishing the construct definition, our next step in scale development was to create an initial pool of items. There are two commonly adopted approaches for generating an item pool: (1) deductive and (2) inductive. The deductive approach implies extracting items from a theoretical perspective based on, for example, a literature review. In contrast, the inductive approach suggests creating items by asking people's viewpoints on a particular subject [34]. Following a mixed process of deductive and inductive approaches has been recommended as a better alternative than using either method in isolation [9, 45] as it overcomes the limitations of using each method independently. Therefore, we used both ways to derive an initial item pool.

To generate an initial pool of items, we performed the following steps:

(1) **Reviewing Existing Literature:** We reviewed existing literature to understand people's experiences and concerns regarding out-of-device threats in the physical world. To the best of our knowledge, we looked for published work where users have reported their concerns or privacy protection practices towards out-of-device threats that match the construct definition, such as shoulder surfing, thermal attacks, smudge attacks and unauthorized access to their device. The reviewed literature included [5, 22, 23, 25, 46, 59]. A total of 20 statements were derived using this approach;
(2) **Item Generation by Researchers:** Two authors (N=2) individually created new items related to the construct definition. A total of 24 statements were developed using this approach;
(3) **Item Generation by a Larger Group of Researchers:** Fellow researchers (N=11) at our institute with expertise in security and privacy and human-computer interaction participated in a short survey. The survey inquired how they would ask people about out-of-device threats in the physical world. A total of 23 statements were constructed using this approach; (Note: A similar item generation approach has been followed by [33, 44].)

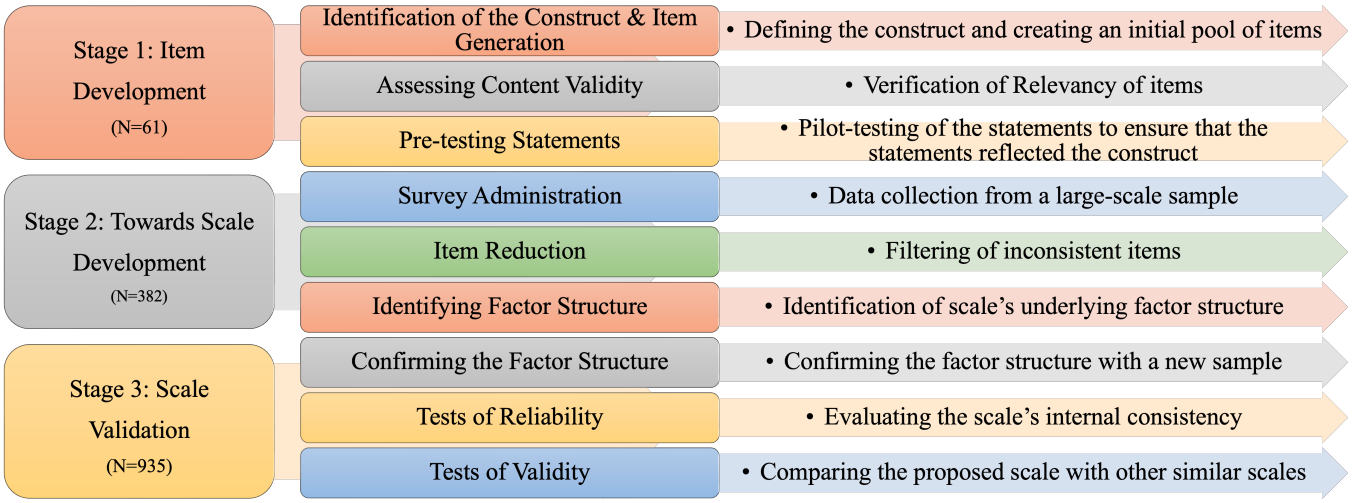

Figure 2: We followed three high-level stages to develop the scale: 1) item development, 2) scale development, and 3) scale validation. At each stage, we followed the recommendations from the literature to refine and develop the scale iteratively. The figure shows the breakdown of the high-level phases carried out in the development of Out-of-Device Privacy Scale (ODPS) along with the sample sizes in each phase.

This procedure of item generation ensured a broad coverage of the construct. The total number of items generated at this point was 67 (the reader is referred to Appendix A for the complete list of items constructed using the various approaches detailed above). This number of items is more than twice the number in the final set (presented in section 5), which fulfils the recommendation by Kline and Schinka et al. [9, 41, 58].

3.3 Refining Items & Assessing Content Validity

Refining items and assessing content validity are crucial in the scale development process and were the next steps after generating an initial pool of items. Moreover, prior work has shown a gap between what privacy scales measure and how they are understood by the general public [17]. Therefore, items should be formulated to exhibit minimal subjective interpretation and be easily understandable by the general public and precise wording [9].

To refine the items and assess content validity, we asked a fellow human-computer interaction (HCI) researcher with a psychology background and English as the native language to review the items and check for four main aspects to account for content validity [65]:

(1) Identification of duplicates or similar worded items,
(2) Verification of item relevance to the construct definition,
(3) Checking of subjective interpretation, and
(4) Inspection of linguistic accuracy.

Out of 67 items, 32 were marked as duplicates, and four were marked as irrelevant to the construct definition. The remaining items were checked and rewritten (if needed) for linguistic accuracy and to minimise subjective interpretation. After the items were reviewed by an HCI expert with a psychology background, two researchers rechecked them and refined them based on the feedback. The items list was reduced to 31 items (see Appendix B).

3.3.1 Pilot Testing. Pilot testing is an essential step before running a study as it helps ensure the smooth running of the study and produces results per the researcher's expectations. Before proceeding with our questionnaire study, we pilot-tested the items from the previous step to check for two things:

(1) **Understanding of Statements:** if the items produced valid measurements based on how easily the public can understand them.
(2) **Variability in Responses:** if the items show wide variability across responses.

The Ethics Committee approved the study at our institute. We deployed the items in an online survey using Qualtrics [53] and advertised it through Prolific [52]. We collected data from 50 participants from the UK. Participants were directed to review item statements concerning privacy and to rate the statements on a 7-point Likert scale based on to what extent they agreed or disagreed with the statement. Additionally, the participants were instructed to identify and report any problems they faced while answering the questions concerning understanding and linguistic accuracy. Keywords, such as "privacy", are often linked to bringing social desirability bias. Researchers warn against the use of such words [21], and the most popular approach to avoid this bias is to use the social desirability scale [19]. However, recent research indicates that the social desirability bias scale does not measure the intended construct [42, 64]. Therefore, we refrained from using the social desirability bias scale. As an alternative, we checked for data skewness and looked for ways to increase data variability by revising the wording of items, which is explained below.

We used attention checks to ensure response accuracy [51] and removed the data of two participants who failed these checks. Among the remaining 48 participants, 13 identified as male, 34

as female, and one as non-binary/third gender. Twenty-nine participants were employed full-time, and eight were employed part-time. Four participants were unemployed, three were retired, two were homemakers, and two were students. Participants aged between 21 years and 68 years (M=37.89, SD=13.56). The median time to complete the questionnaire was 6.32 minutes. Participants were compensated with 1.10 USD, following the set standard by Prolific for their time.

Most of the statements were easily understood by participants, and participants did not report any significant issues. Participants mentioned a few problems regarding some item statements; for example, "I use biometric authentication to avoid someone observing my password and/or to avoid any oily or heat residues on the screen." was not clear as to which biometric authentication is being referred. Based on the feedback from the participants, we revised the wording of several item statements. Next, we checked for variability by observing descriptive statistics and noted that wide variability was found across most items. No items were removed based on their variability. Finally, two researchers reviewed the complete statements again and inspected for issues or similar wording. A few item statements were removed as they were very similarly worded to other statements. After this step, 26 item statements were retained for further analysis (see Appendix C).

4 STAGE 2: TOWARDS DEVELOPING THE SCALE

Towards developing our out-of-device privacy scale, the next step after item generation was developing the scale. This section explains the data collection process, the participants' details, our data checks, and the results of the exploratory factor analysis.

4.1 Initial Data Analysis

A sample of N=400 participants was recruited on Prolific to investigate our 26-item questionnaire from the previous step. The sample size was determined following recommendations Nunnally [49], which shows there should be at least ten responses per statement. Further, ten responses per item are among the best practices for scale development [67]. Participants who participated in the pilot study were excluded from taking part in this study. All questions were randomized to avoid order effects.

Eighteen participants were removed from the analysis as they failed the attention check. Out of the remaining N=382 participants, N=135 identified as male, N=242 as female, N=3 preferred not to disclose, and N=2 self-described as third gender/non-binary. Participants were, on average, 41.16 years old (SD=13.5, Min=18, Max=83). N=214 participants were full-time employed, whereas N=86 were part-time employed. Twenty participants were unemployed, N=14 students, N=17 homemakers, and N=31 retired. All participants were based in the UK and were compensated for participation by the Prolific recommendation.

Before proceeding with the factor analysis, we checked the data for variability, which included checking for descriptive statistics. The means of the statements were between 3.5 and 5.9, except for two items, 2.13 and 2.104. The SDs were between 1 and 2. Medians ranged from 3.5 to 6 except for two items: (1) *"To avoid people nearby from looking at my smartphone screen, I specifically use a privacy-protecting screen cover (e.g., tampered glass protector)"*, and (2) *"I press extra keys after I have entered my PIN at the atm to avoid anyone taking a heat-trace picture of my PIN"*. No items were removed based on their response distribution.

Next, we calculated the item-total correlation of items and four items were removed as their item-total correlation was less than 0.30 [9]. We then checked for internal consistency amongst the remaining 22 items using Cronbach's alpha [18]. The items exhibited high internal consistency (α=0.884).

4.2 Exploratory Factor Analysis

As a pre-requisite to establishing the number of factors and their structure, we first evaluated the suitability of our dataset, whether it measured common factors, and whether they were correlated. We checked this by performing the Kaiser-Meyer-Olkin (KMO) Measures of Sampling Adequacy (MSA) test [14]. KMO indicates how much a correlation matrix contains factors or simple chance correlations. A KMO value of 0.60 or higher is appropriate for factor analysis [65]. In our case, the entire dataset had a KMO value of 0.914, considered "marvellous" [38] and thus well within the bounds of adequacy. None of the items had a lower KMO (*i.e.*, less than 0.5). Bartlett's test of sphericity (x^2 = 2561.179, ρ<.001) further confirmed that the set of items was suitable for factor analysis [63, 67].

Next, the visual inspection of the scree plot revealed that the kink was between two and three factors, as seen in Figure 3. The "kink" in the scree plot indicates the number of factors we should be looking for [11]. Using the elbow method, the scree plot suggested a single-factor solution. To confirm the interpretation, we explored two and three-factor solutions. We performed principal axis factoring with varimax rotation using a loading cut-off of 0.35, the recommended threshold [9, 30]. The three-factor solution did not give a meaningful output, and two of the factors had a conceptual overlap. Therefore, we next explored a two-factor solution. The two-factor solution produced a simple structure. To decide between the two-factor solution and a single single-factor solution, we checked for correlations between the factors in the 2-factor solution and calculated the Pearson correlation.

For this purpose, scores of each factor were calculated by averaging the constituent items. There was a statistically significant positive correlation between the two factors (r=0.635, p<.001). The high correlation between the factors, the scree plot interpretation and the above discussion suggested a single-factor solution would be suitable. To further confirm this, we explored multi-factor solutions using an oblique rotation (i.e. direct oblimin) as detailed in Appendix E. This further analysis confirmed that a single-factor solution is more suitable.

After deciding on the single-factor solution, we proceeded with further analysis. Four items loaded below 0.40, the recommended minimum threshold of loading [65]. We removed three out of four items for this reason. Still, we kept the item (*item 5*) with a loading of 0.371 as we felt this item represented a particular attribute of out-of-device privacy, i.e. concern for other people's privacy, and was not captured elsewhere in the set of item statements. Further, the loading of this particular item was not lower than 0.35 (unlike the remaining), which is the minimum threshold and was close to 0.4.

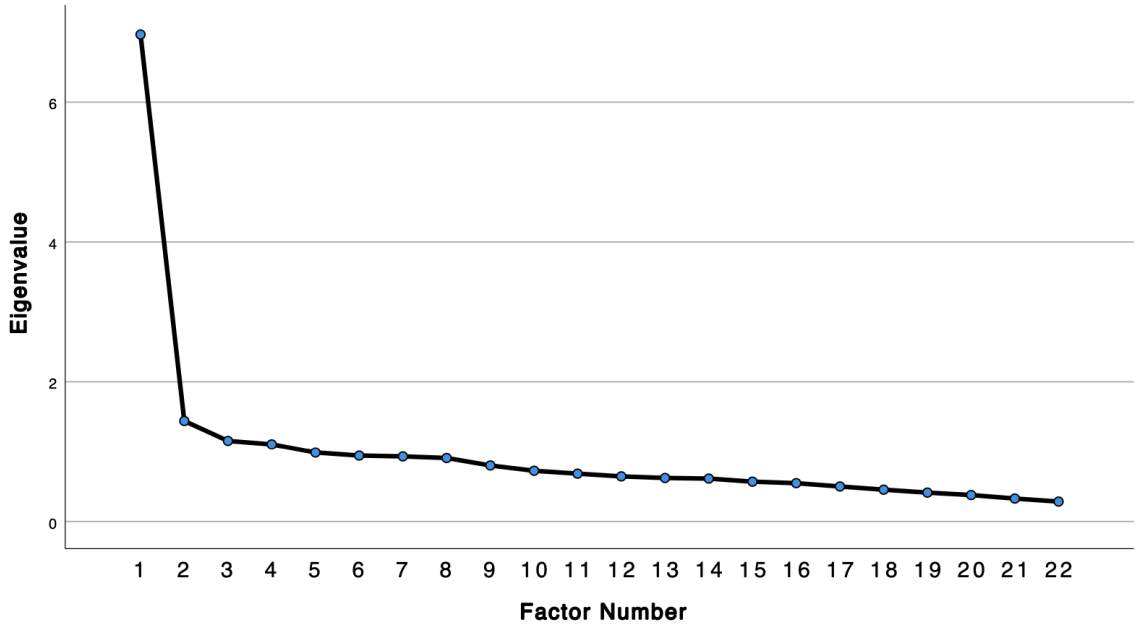

Figure 3: The figure shows the scree plot produced using the data for Exploratory Factor Analysis to determine the appropriate number of factors.

IN	Item Statement	FL
21	It is important for me to protect my screen content from people around me on public transport	0.77
25	I am worried that someone might access my information by spying on what I am doing on my smartphone	0.684
26	I get anxious when someone from my surroundings invades my privacy by looking at the screen	0.711
24	I am a privacy-centred person	0.643
17	I consider my data to be a target for external threats to my device such as shoulder surfing	0.609
18	I believe my data is worth protecting from external threats to my device	0.452
20	It is not fine for me to have my smartphone screen visible to the public	0.591
23	I mind if a stranger sitting next to me takes a look at my smartphone while I am watching a private video	0.512
19	Privacy invasion by people surrounding us is effective in leaking information	0.513
14	The increasing availability and affordability of audio, video and photo recording devices are a threat to everyone's privacy	0.449
9	I feel concerned when using ATMs that use cameras for recording purposes	0.456
15	I would like my device to do something to alert me every time someone looks at it without my permission	0.429
8	I would change all of my passwords immediately if my smartphone was lost	0.451
5	If someone sees my friend's content on my screen, it feels like a breach of my friend's trust in me to keep their content private	0.371
12	Among the reasons I use auto-fill for passwords on my smartphone, is to avoid anyone overseeing what I enter	0.477
16	I scroll quickly when I sense someone is looking at my smartphone's screen	0.606
1	I try to adjust my hand position when using my smartphone so that no one can see the information shown on it	0.557
13	I check for any surrounding people when I am doing something on my smartphone in public places	0.644
2	If anyone looks at my screen without permission, I tend to put my smartphone away	0.576
4	I lower my screen brightness so that no one around me can take a look at what is shown on it	0.337
3	To avoid people nearby from looking at my smartphone screen, I specifically use a privacy-protecting screen cover (e.g. tampered glass protector)	0.306
6	I use fingerprint (or other biometric methods) mainly to avoid someone observing my password	0.306

Table 1: The Table shows the results of Exploratory Factor Analysis. The items marked in red (last three items) were removed from further analysis as they did not load sufficiently high. IN = Item Number; FL = Factor Loading.

For this reason, *item 5* was included in the scale questionnaire. The three removed items are marked in red in Table 1. Next, we checked for Cronbach's alpha after removing the three items: (α=0.888). The 19 items from this step were retained for further analysis in the following steps.

5 STAGE 3: FINAL SCALE VALIDATION

We performed a confirmatory factor analysis to confirm the previous section's explored factor structure. For this, we deployed the 19-item questionnaire using Qualtrics as an online questionnaire on Prolific. Participants from previous studies (pilot test and study 1) were excluded from participation, ensuring a new independent sample [9].

5.1 Study Design

The online questionnaire included the 19 items shortlisted from the previous section. In the questionnaire, we also had the items from the IUIPC [44] and CFIP [62], which measure online privacy concerns to investigate how closely our proposed scale is related to other privacy constructs[1] All items from IUIPC and CFIP used a 7-point Likert scale, and responses ranged from "strongly disagree" to "strongly agree". All questions were randomized to avoid sequence effects [55]. We also used attention check questions for participants' attention and responsiveness towards the study [51].

5.2 Participants

We collected data from 1,000 participants, out of which N=65 failed the attention check, and therefore, their data was excluded from the analysis. From the remaining N=935, N=554 self-identified as females, N=371 as males, N=8 as non-binary/third gender, and N=2 preferred not to say. Participants were aged between 19 and 90 (Mean=43.39, SD=13.81). Most participants were employed full-time (N=515), while N=179 were employed part-time. N=83 participants were retired, N=68 were unemployed, N=52 were homemakers, and N=38 were students. All participants were based in the UK and were compensated for their time per the Prolific's recommended rate.

5.3 Initial Data Analysis

Before proceeding with the confirmatory factor analysis, we checked the data's suitability by computing the Kaiser-Meyer-Olkin (KMO) Measures of Sampling Adequacy (MSA) test. The full dataset has KMO = 0.956, and none of the items had a lower KMO than 0.907 [67]. Bartlett's test of sphericity (x^2 = 7562.784, ρ<.0005) further confirmed that the set of items was suitable for factor analysis [63].

5.4 Confirmatory Factor Analysis

We performed a confirmatory factor analysis to provide statistical support to the explored factor structure in the previous section. We calculated the following fit indices: the Root Mean Square Error of Approximation (RMSEA), the Comparative Fit Index (CFI), and the Tucker-Lewis Index (TLI) [9]. We intentionally did not consider chi-square goodness-of-fit as it is reported to be unreliable for large data samples [35, 60]. The results revealed that CFI = 0.911, TLI = 0.889, and RMSEA = 0.068. While RMSEA and CFI indicated acceptable fit, TLI was slightly lower than the recommendation (i.e. should be greater than 0.9) [48].

We followed a step-wise model selection procedure with backward elimination to improve TLI by checking for item loadings. The item with the lowest loading was removed to see if it improved TLI. The item with the least loading was *item 9*. We found that by eliminating *item 9*, all indices resulted in a good model fit (i.e. CFI=0.923, TLI= 0.903, RMSEA=0.066) [7, 40, 43, 48]. We assume that this may be because only *item 9* related to using ATMs. In contrast, no other items were associated with or about ATMs. Thus, our final scale contains 18 items, as shown in Table 3.

[1]Please, note: We are aware that the CFIP "collection" subscale is repeated in the 10-item IUIPC scale, but we included it anyway and made the comparison as both subscales differ slightly in items' wording. We report the results with both subscale "collection" versions.

5.5 Tests of Reliability

Reliability is the internal consistency commonly measured using Cronbach's alpha [11, 18]. A coefficient of 0.70 or higher is considered acceptable. For the 18 items, Cronbach's alpha was 0.917, well above the recommendation of 0.70 [9]. Next, we computed the composite reliability score [54], which turned out to be 0.922, high above the recommended threshold of 0.60 [4].

5.6 Construct Validity

Scale validity refers to whether the measured concept fully corresponds to the construct it aims to measure[11]. This defines construct validity, which is the foundation of any questionnaire [12]. We performed a convergent validity analysis to assess the construct validity of ODPS.

We compared ODPS with the 10-item IUIPC and CFIP to assess convergent validity. We hypothesized a positive relation between the subscales of IUIPC and CFIP overall, as all three scales relate to privacy but capture different dimensions of privacy. Table 2 shows the results of the correlations and percentage variability along with the reliability score calculated using Cronbach's alpha [18] for each of the subscales of IUIPC and CFIP. Our scale demonstrated positive correlations with all subscales of IUIPC and CFIP (p<0.001). However, none of the correlations exceeds 14.44% variability, showing a maximum of 14.44% conceptual overlap of ODPS with the compared subscales. This much overlap is expected as online privacy and out-of-device privacy, both fall under the privacy umbrella. However, 85.56% total variability cannot be explained by concerns about how organizations handle data privacy or online privacy concerns. Therefore, we conclude that our scale (ODPS), which measures out-of-device privacy, differs from how organizations collect, process, store, and use information (i.e. IUIPC and CFIP).

		τ	Variability %	α
IUIPC	Control	0.368	13.54	0.727
	Awareness	0.38	14.22	0.732
	Collection	0.359	12.89	0.904
CFIP	Errors	0.303	9.18	0.86
	Unauthorized Use	0.222	4.93	0.827
	Improper Access	0.304	9.24	0.804
	Collection	0.327	10.69	0.93
	Overall CFIP	0.377	14.21	0.898

Table 2: The Table shows (1) the correlation (Kendall's Tau) between ODPS and IUIPC & CFIP and (2) the reliability score of each of the subscales in our dataset of the second study.

6 DISCUSSION

In this paper, we contribute a reliable and valid psychometric instrument to measure the out-of-device privacy of users that describes the importance a person attributes to protecting personal information from threats out of the device in the physical world. We detail the rigorous methodology adapted to develop and refine the scale

questionnaire. The 18-item scale fills the gap in protecting against out-of-device threats in the physical world.

6.1 Obstructions in Scale Development Studies

Prior work has identified two key obstructions in scale development studies: 1) understanding of scale statements by the general public and 2) verifying if the scale measures the construct it aims to measure [17]. Assessing and ensuring these two key points are crucial in the scale development study as they provide accurate measurements. In our research, we took extra care to ensure both key points were accessed and checkmarked. For example, before beginning the factor analysis, we confirmed whether the general public understood the scale statements accurately. For this, the statements were checked and revised by an HCI expert with a psychology background for relevancy to the construct definition and subjective interpretation. Two researchers then rechecked the items to double-ensure the results. The items were then pilot-tested with a small sample from the general public to check for understanding (see Section 3.3). Further, after finalising the scale statements through factor analysis and confirmatory factor analysis, we performed tests for convergent validity to ensure that the scale fully corresponds to the construct it aims to measure (see Section 5.6). Therefore, we conclude that we have confidently assessed and ensured high-quality scale development while eliminating the obstacles.

6.2 Using ODPS to Measure Out-of-Device Privacy

ODPS will be helpful to researchers who aim to mitigate privacy threats in the physical world. It can be easily deployed in an online questionnaire format and distributed on a large scale. ODPS would provide insights into the user privacy profile, which could then be used to inform the design of protection mechanisms. Further, researchers can utilize it to measure the privacy behaviour of a user group. The scale can help explore how privacy perception changes over time and across different user groups. ODPS can offer to answer research questions like: *To what degree are users concerned about protecting their information from privacy threats in the physical world?* or *What is the users' level of awareness of privacy threats in the physical world?* or *How much users are willing to do to protect their privacy from threats in the physical world?*.

6.3 Using ODPS to design Protection Mechanisms

ODPS will be advantageous in designing protection methods against threats in the physical world, such as shoulder surfing. For example, a low ODPS score would indicate that the user prefers a light protection method. In contrast, a high ODPS score would mean the user is highly concerned about privacy and prefers a strong protection method. In the same way, the ODPS score could reflect user awareness of privacy threats in the physical world. This could empower users to defend their privacy with and without a device-based mechanism. Researchers could develop awareness strategies based on the scores to educate users on privacy violations.

Further, ODPS could also be used to enhance the design of protection mechanisms by controlling participants' out-of-device privacy attitudes. For example, in a usability or user experience evaluation study of a protection mechanism, ODPS could serve as a covariate to ensure that the participant's experience outcome is the result of the change in the design of the protection mechanism and not due to the differences in their out-of-device privacy. In summary, ODPS could offer to investigate research questions like: *What is users' level of awareness of privacy threats in the physical world?* or *How can the design of protection mechanisms improve to reflect the user's out-of-device privacy perception better?*.

6.4 Using ODPS to measure Privacy Culture

The perception of privacy changes as we move across cultures [57]. For example, shoulder surfing can have severe consequences in some cultures like the Middle East, whereas it is sometimes ignored in Western cultures [22, 57]. ODPS would help measure the privacy culture, which could be incorporated into the tech devices. Based on this, users can be offered personalized protection based on their cultural setting.

6.5 Using ODPS in Combination with Other Privacy Scales

ODPS, in combination with other privacy scales such as IUIPC [44], can help construct a privacy profile of users which would explain users' perceptions of privacy in the online and physical world, summing up as a complete privacy profile. This privacy profile could then be used to provide holistic protection to user's information online and in the physical world.

6.6 Instructions for Scoring

ODPS is a psychometric instrument developed to measure the out-of-device privacy of users. To use the scale questionnaire, the statements should be presented using a Likert scale with 7 points, starting from strongly disagree to strongly agree. All items are mandatory to answer, and no item requires reverse scoring. The scale statements should be randomized to avoid order effects. The scale score can be calculated by averaging the components' scores.

6.7 Limitations & Future Work

Scale validation is a continuous process. While we followed the best practices from the literature in iteratively developing and refining the scale, further studies must be conducted to provide statistical strength to ODPS. Second, while we recruited a large number of participants, all participants were based in the United Kingdom. This might have introduced selection bias. Further studies with participants from different geographic locations should be conducted to strengthen the validation of the scale. Third, while we followed the most recommended approach for item generation and item elimination, it may be possible that some possible factors were not captured during our process. Although our items produce reliable and valid results, future work should expand on the findings. Lastly, we propose that in future studies, the scale should be administered in mechanisms developer studies to investigate ODPS's impact on the design of privacy protection mechanisms.

	Item Statement
1	It is important for me to protect my screen content from people around me on public transport
2	I am worried that someone might access my information by spying on what I am doing on my smartphone
3	I get anxious when someone from my surroundings invades my privacy by looking at the screen
4	I am a privacy-centred person
5	I consider my data to be a target for external threats to my device such as shoulder surfing
6	I believe my data is worth protecting from external threats to my device
7	It is not fine for me to have my smartphone screen visible to the public
8	I mind if a stranger sitting next to me takes a look at my smartphone while I am watching a private video
9	Privacy invasion by people surrounding us is effective in leaking information
10	The increasing availability and affordability of audio, video and photo recording devices are a threat to everyone's privacy
11	I would like my device to do something to alert me every time someone looks at it without my permission
12	I would change all of my passwords immediately if my smartphone was lost
13	If someone sees my friend's content on my screen, it feels like a breach of my friend's trust in me to keep their content private
14	Among the reasons I use auto-fill for passwords on my smartphone, is to avoid anyone overseeing what I enter
15	I scroll quickly when I sense someone is looking at my smartphone's screen
16	I try to adjust my hand position when using my smartphone so that no one can see the information shown on it
17	I check for any surrounding people when I am doing something on my smartphone in public places
18	If anyone looks at my screen without permission, I tend to put my smartphone away

Table 3: The table shows the final look of the 18-item Out-of-Device Privacy Scale (ODPS)

7 CONCLUSION

In this paper, we present a reliable and valid 18-item psychometric scale, the "out-of-device Privacy Scale (ODPS)", to capture the out-of-device privacy of users. We followed the best scale development practices from the literature, ensuring a rigorous methodology. We present a detailed description of each step of the development and validation of the scale. With the aid of ODPS, privacy and security researchers will be assisted in designing user-centred protection mechanisms offering personalized and holistic protection against out-of-device threats in the physical world.

ACKNOWLEDGMENTS

We thank all participants for their time and participation. We are also grateful to Prof Paul Cairns (University of York), Dr Graham Wilson and Dr Shaun Macdonald (University of Glasgow) for their support and guidance throughout the project. This publication was supported by an Excellence Bursary Award by the University of Glasgow, the Scottish Informatics & Computer Science Alliance (SICSA), an EPSRC New Investigator Award (grant number EP/V008870/1), and the PETRAS National Centre of Excellence for IoT Systems Cybersecurity, which is also funded by the UK EPSRC under grant number EP/S035362/1. This work was supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy - EXC 2092 CASA - 390781972. Lastly, we would like to acknowledge CANVA for offering support in the form of a Free Content License (Figure 1 was created using Canva [13] under Free Content License).

REFERENCES

[1] Yomna Abdelrahman, Mohamed Khamis, Stefan Schneegass, and Florian Alt. 2017. Stay cool! understanding thermal attacks on mobile-based user authentication. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. 3751–3763.

[2] Norah Alotaibi, John Williamson, and Mohamed Khamis. 2023. ThermoSecure: Investigating the effectiveness of AI-driven thermal attacks on commonly used computer keyboards. *ACM Transactions on Privacy and Security* 26, 2 (2023), 1–24.

[3] Amazon. 2023. *Amazon Listing - PerfectPrime IR203, (IR) Infrared Thermal Imager Camera*. Accessed: 2023-01-13.

[4] Richard P Bagozzi and Youjae Yi. 1988. On the evaluation of structural equation models. *Journal of the academy of marketing science* 16 (1988), 74–94.

[5] Paul Bekaert, Norah Alotaibi, Florian Mathis, Nina Gerber, Aidan Christopher Rafferty, Mohamed Khamis, and Karola Marky. 2022. Are thermal attacks a realistic threat? Investigating the preconditions of thermal attacks in users' daily lives. In *Nordic Human-Computer Interaction Conference*. 1–9.

[6] France Bélanger and Robert E Crossler. 2011. Privacy in the digital age: a review of information privacy research in information systems. *MIS quarterly* (2011), 1017–1041.

[7] Peter M Bentler. 1990. Comparative fit indexes in structural models. *Psychological bulletin* 107, 2 (1990), 238.

[8] Antonio Bianchi, Jacopo Corbetta, Luca Invernizzi, Yanick Fratantonio, Christopher Kruegel, and Giovanni Vigna. 2015. What the app is that? deception and countermeasures in the android user interface. In *2015 IEEE Symposium on Security and Privacy*. IEEE, 931–948.

[9] Godfred O Boateng, Torsten B Neilands, Edward A Frongillo, Hugo R Melgar-Quiñonez, and Sera L Young. 2018. Best practices for developing and validating scales for health, social, and behavioral research: a primer. *Frontiers in public health* 6 (2018), 149.

[10] Tom Buchanan, Carina Paine, Adam N Joinson, and Ulf-Dietrich Reips. 2007. Development of measures of online privacy concern and protection for use on the Internet. *Journal of the American society for information science and technology* 58, 2 (2007), 157–165.

[11] Paul Cairns. 2019. *Doing better statistics in human-computer interaction*. Cambridge University Press.

[12] Paul Cairns, M Soegaard, and RF Dam. 2016. Experimental methods in human-computer interaction. *Encyclopedia of Human-Computer Interaction* (2016).

[13] Canva. 2022. Canva. https://www.canva.com

[14] Barbara A Cerny and Henry F Kaiser. 1977. A study of a measure of sampling adequacy for factor-analytic correlation matrices. *Multivariate behavioral research* 12, 1 (1977), 43–47.

[15] Ankur Chattopadhyay and Terrance E Boult. 2007. Privacycam: a privacy preserving camera using uclinux on the blackfin dsp. In *2007 IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, 1–8.

[16] Mark H Chignell, Anabel Quan-Haase, and Jacek Gwizdka. 2003. The privacy attitudes questionnaire (paq): initial development and validation. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, Vol. 47. SAGE Publications Sage CA: Los Angeles, CA, 1326–1330.

[17] Jessica Colnago, Lorrie Faith Cranor, Alessandro Acquisti, and Kate Hazel Stanton. 2022. Is it a concern or a preference? An investigation into the ability of privacy scales to capture and distinguish granular privacy constructs. In *Eighteenth Symposium on Usable Privacy and Security (SOUPS 2022)*. 331–346.

[18] Lee J Cronbach. 1950. Further evidence on response sets and test design. *Educational and psychological measurement* 10, 1 (1950), 3–31.

[19] Douglas P Crowne and David Marlowe. 1960. A new scale of social desirability independent of psychopathology. *Journal of consulting psychology* 24, 4 (1960), 349.

[20] Youngwook Do, Nivedita Arora, Ali Mirzazadeh, Injoo Moon, Eryue Xu, Zhihan Zhang, Gregory D. Abowd, and Sauvik Das. 2023. Powering for Privacy: Improving User Trust in Smart Speaker Microphones with Intentional Powering and Perceptible Assurance. In *32nd USENIX Security Symposium (USENIX Security 23)*. USENIX Association, Anaheim, CA, 2473–2490. https://www.usenix.org/

conference/usenixsecurity23/presentation/do
[21] Serge Egelman and Eyal Peer. 2015. Scaling the security wall: Developing a security behavior intentions scale (sebis). In *Proceedings of the 33rd annual ACM conference on human factors in computing systems*. 2873–2882.
[22] Malin Eiband, Mohamed Khamis, Emanuel Von Zezschwitz, Heinrich Hussmann, and Florian Alt. 2017. Understanding shoulder surfing in the wild: Stories from users and observers. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. 4254–4265.
[23] Habiba Farzand, Kinshuk Bhardwaj, Karola Marky, and Mohamed Khamis. 2021. The Interplay between Personal Relationships & Shoulder Surfing Mitigation. In *Mensch und Computer 2021*. 338–343.
[24] Habiba Farzand, Karola Marky, and Mohamed Khamis. 2022. I Hate When People Do This; There'sa Lot of Sensitive Content for Me": A Typology of Perceived Privacy-Sensitive Content in Shoulder Surfing Scenarios. In *Proceedings of the Eighteenth USENIX Conference on Usable Privacy and Security. USENIX Association, USA*.
[25] Habiba Farzand, Karola Marky, and Mohamed Khamis. 2022. Shoulder Surfing through the Social Lens: A Longitudinal Investigation & Insights from an Exploratory Diary Study. In *Proceedings of the 2022 European Symposium on Usable Security*. 85–97.
[26] Habiba Farzand, Karola Marky, and Mohamed Khamis. 2023. "... It's very unacceptable for someone to peek into your privacy." Chronicles of Shoulder Surfing: Exploring Deep into a Longitudinal Diary Study. In *Proceedings of the Nineteenth USENIX Conference on Usable Privacy and Security. USENIX Association, USA*.
[27] Habiba Farzand, Karola Marky, and Mohamed Khamis. 2023. "... It's very unacceptable for someone to peek into your privacy." Chronicles of Shoulder Surfing: Exploring Deep into a Longitudinal Diary Study. In *Proceedings of the Nineteenth USENIX Conference on Usable Privacy and Security. USENIX Association, USA*.
[28] Nina Gerber, Verena Zimmermann, and Melanie Volkamer. 2019. Why johnny fails to protect his privacy. In *2019 IEEE European Symposium on Security and Privacy Workshops (EuroS&PW)*. IEEE, 109–118.
[29] Wendy Goucher. 2011. Look behind you: the dangers of shoulder surfing. *Computer Fraud & Security* 2011, 11 (2011), 17–20.
[30] Joseph F Hair, RE Anderson, RL Tatham, and WC Black. 1998. Multivariate data analysis prentice hall. *Upper Saddle River, NJ* 730 (1998).
[31] Marian Harbach, Emanuel Von Zezschwitz, Andreas Fichtner, Alexander De Luca, and Matthew Smith. 2014. It's a hard lock life: A field study of smartphone (un) locking behavior and risk perception. In *10th Symposium On Usable Privacy and Security ({SOUPS} 2014)*. 213–230.
[32] Marian Harbach, Emanuel Von Zezschwitz, Andreas Fichtner, Alexander De Luca, and Matthew Smith. 2014. {It's} a hard lock life: A field study of smartphone ({Un) Locking} behavior and risk perception. In *10th symposium on usable privacy and security (SOUPS 2014)*. 213–230.
[33] Rakibul Hasan, Rebecca Weil, Rudolf Siegel, and Katharina Krombholz. 2023. A Psychometric Scale to Measure Individuals' Value of Other People's Privacy (VOPP). In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–14.
[34] Timothy R Hinkin. 1998. A brief tutorial on the development of measures for use in survey questionnaires. *Organizational research methods* 1, 1 (1998), 104–121.
[35] Li-tze Hu and Peter M Bentler. 1999. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural equation modeling: a multidisciplinary journal* 6, 1 (1999), 1–55.
[36] Harris Interactive. 2002. Privacy on and off the Internet: What consumers want. *Privacy and American Business* (2002), 1–127.
[37] Lijun Jiang, Weizhi Meng, Yu Wang, Chunhua Su, and Jin Li. 2017. Exploring energy consumption of juice filming charging attack on smartphones: a pilot study. In *International Conference on Network and System Security*. Springer, 199–213.
[38] Henry F Kaiser. 1974. An index of factorial simplicity. *psychometrika* 39, 1 (1974), 31–36.
[39] Mohamed Khamis, Malin Eiband, Martin Zürn, and Heinrich Hussmann. 2018. EyeSpot: Leveraging Gaze to Protect Private Text Content on Mobile Devices from Shoulder Surfing. *Multimodal Technologies and Interaction* 2, 3 (2018), 45.
[40] Hyunho Kim, Boncho Ku, Jong Yeol Kim, Young-Jae Park, Young-Bae Park, et al. 2016. Confirmatory and exploratory factor analysis for validating the phlegm pattern questionnaire for healthy subjects. *Evidence-Based Complementary and Alternative Medicine* 2016 (2016).
[41] Paul Kline. 2013. *Handbook of psychological testing*. Routledge.
[42] Lukas Lanz, Isabel Thielmann, and Fabiola H Gerpott. 2022. Are social desirability scales desirable? A meta-analytic test of the validity of social desirability scales in the context of prosocial behavior. *Journal of Personality* 90, 2 (2022), 203–221.
[43] Robert C MacCallum, Michael W Browne, and Hazuki M Sugawara. 1996. Power analysis and determination of sample size for covariance structure modeling. *Psychological methods* 1, 2 (1996), 130.
[44] Naresh K Malhotra, Sung S Kim, and James Agarwal. 2004. Internet users' information privacy concerns (IUIPC): The construct, the scale, and a causal model. *Information systems research* 15, 4 (2004), 336–355.
[45] Fabiane FR Morgado, Juliana FF Meireles, Clara M Neves, Ana Amaral, and Maria EC Ferreira. 2017. Scale development: ten main limitations and recommendations to improve future research practices. *Psicologia: Reflexão e Crítica* 30 (2017).
[46] Ildar Muslukhov, Yazan Boshmaf, Cynthia Kuo, Jonathan Lester, and Konstantin Beznosov. 2012. Understanding users' requirements for data protection in smartphones. In *2012 IEEE 28th international conference on data engineering workshops*. IEEE, 228–235.
[47] Ildar Muslukhov, Yazan Boshmaf, Cynthia Kuo, Jonathan Lester, and Konstantin Beznosov. 2013. Know your enemy: the risk of unauthorized access in smartphones by insiders. In *Proceedings of the 15th international conference on Human-computer interaction with mobile devices and services*. 271–280.
[48] Richard G Netemeyer, William O Bearden, and Subhash Sharma. 2003. *Scaling procedures: Issues and applications*. sage publications.
[49] Jum C Nunnally. 1967. *Psychometric theory*. McGraw-hill.
[50] Maggie Oates, Yama Ahmadullah, Abigail Marsh, Chelse Swoopes, Shikun Zhang, Rebecca Balebako, and Lorrie Faith Cranor. 2018. Turtles, locks, and bathrooms: Understanding mental models of privacy through illustration. *Proceedings on Privacy Enhancing Technologies* 2018, 4 (2018), 5–32.
[51] Daniel M Oppenheimer, Tom Meyvis, and Nicolas Davidenko. 2009. Instructional manipulation checks: Detecting satisficing to increase statistical power. *Journal of experimental social psychology* 45, 4 (2009), 867–872.
[52] Prolific. 2023. Prolific | Online participant recruitment for surveys and market research. https://www.prolific.co/ Retrieved September 01, 2023.
[53] Qualtrics. 2021. Qualtrics - Leading Experience Management and Survey Software. https://www.qualtrics.com/uk/?rid=ip&prevsite=en&newsite=uk&geo=GB&geomatch=uk Retrieved February 11, 2021.
[54] Tenko Raykov. 1997. Estimation of composite reliability for congeneric measures. *Applied Psychological Measurement* 21, 2 (1997), 173–184.
[55] Ulf-Dietrich Reips. 2002. Standards for Internet-based experimenting. *Experimental psychology* 49, 4 (2002), 243.
[56] Alia Saad, Michael Chukwu, and Stefan Schneegass. 2018. Communicating Shoulder Surfing Attacks to Users. In *Proceedings of the 17th International Conference on Mobile and Ubiquitous Multimedia*. 147–152.
[57] Mennatallah Saleh, Mohamed Khamis, and Christian Sturm. 2019. What About My Privacy, Habibi?. In *IFIP Conference on Human-Computer Interaction*. Springer, 67–87.
[58] John A Schinka and Wayne F Velicer. 2003. *Handbook of Psychology, Developmental Psychology*. Vol. 6. John Wiley & Sons.
[59] Stefan Schneegass, Frank Steimle, Andreas Bulling, Florian Alt, and Albrecht Schmidt. 2014. Smudgesafe: Geometric image transformations for smudge-resistant user authentication. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. 775–786.
[60] Dexin Shi, Christine DiStefano, Heather L McDaniel, and Zhehan Jiang. 2018. Examining chi-square test statistics under conditions of large model size and ordinal data. *Structural Equation Modeling: A Multidisciplinary Journal* 25, 6 (2018), 924–945.
[61] Christopher Slobogin. 2002. Public privacy: camera surveillance of public places and the right to anonymity. *Miss. LJ* 72 (2002), 213.
[62] H Jeff Smith, Sandra J Milberg, and Sandra J Burke. 1996. Information privacy: Measuring individuals' concerns about organizational practices. *MIS quarterly* (1996), 167–196.
[63] George W Snedecor and Witiiam G Cochran. 1989. Statistical methods, 8thEdn. *Ames: Iowa State Univ. Press Iowa* 54 (1989), 71–82.
[64] Alvaro Vergés. 2022. On the desirability of social desirability measures in substance use research. *Journal of Studies on Alcohol and Drugs* 83, 4 (2022), 582–587.
[65] Roger L Worthington and Tiffany A Whittaker. 2006. Scale development research: A content analysis and recommendations for best practices. *The counseling psychologist* 34, 6 (2006), 806–838.
[66] Guixin Ye, Zhanyong Tang, Dingyi Fang, Xiaojiang Chen, Kwang In Kim, Ben Taylor, and Zheng Wang. 2017. Cracking android pattern lock in five attempts. In *Proceedings of the 2017 Network and Distributed System Security Symposium 2017 (NDSS 17)*. Internet Society.
[67] An Gie Yong, Sean Pearce, et al. 2013. A beginner's guide to factor analysis: Focusing on exploratory factor analysis. *Tutorials in quantitative methods for psychology* 9, 2 (2013), 79–94.
[68] Huiyuan Zhou, Khalid Tearo, Aniruddha Waje, Elham Alghamdi, Thamara Alves, Vinicius Ferreira, Kirstie Hawkey, and Derek Reilly. 2016. Enhancing Mobile Content Privacy with Proxemics Aware Notifications and Protection. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. 1362–1373.

A ITEM GENERATION PHASE

This section lists the total items created in the initial item generation phase using literature-based and empirical approaches.

A.1 Items Created Using Literature-Based Approach

(1) I turn off my device's display if I notice someone looking over my screen without my permission[22]
(2) I adjust my position when browsing through my smartphone so that no one can take a look at it [23]
(3) I hide the screen with my hands if I notice someone looking at my screen [25]
(4) I usually avoid accessing apps that contain private information when I am around others [22]
(5) If someone looks at my screen without permission, I usually ignore them [22]
(6) To avoid surrounding people from looking at my screen, I use a tampered privacy protector on my device [23]
(7) I lower my screen brightness so that no one around me can take a look at my screen [25]
(8) I am concerned about my reputation if someone oversees my device screen without my permission [23]
(9) If someone oversees my screen content, I would feel uncomfortable, because I feel like those people trusted me to keep their data private [22]
(10) I often clean my device screen to remove any oily residues so that no one can use them to trace what I entered on the device [59]
(11) I use biometric authentication to avoid someone observing my password and/or to avoid any oily or heat residues on the screen [5]
(12) I backup my valuable data often for safety in case of theft or lost [46]
(13) I carry a small paper book to save my contacts in case my device is lost/stolen [46]
(14) I back up my data frequently [46]
(15) I do not trust the security of smartphones and therefore do not store any sensitive information on them [46]
(16) I would change all my passwords immediately in my smartphone is stolen or lost [46]
(17) I do not leave my device unattended around others to avoid giving anyone the opportunity to unlock it [5]
(18) I am concerned when using ATMs that use cameras for recording purposes [5]
(19) I wear gloves to avoid anyone taking heat traces picture of my PIN when I use ATM [5]
(20) I press extra keys after I have entered my PIN at the ATM [5]

A.2 Items Constructed Through Deductive Approach - Researchers Developed Items

(21) I clear my location track history because in case my phone is lost then someone might be able to track down my home
(22) I access sensitive data only on personal PC
(23) I keep my security knowledge up to date
(24) I use auto-fill in passwords to avoid anyone over seeing my passwords when I enter on my smartphone
(25) I am concerned about my reputation if I see someone looking over my screen without permission
(26) I check for my surrounding people when I use my smartphone at public places
(27) The survellance cameras concern me as I fear that they might be recording my device interaction
(28) The increasing availability and cost feasibility of devices like thermal cameras are a threat to everyone's privacy
(29) I clean my smartphone screen often to clear off any smudges left behind after interaction
(30) I use two smartphones, one for private and indoor usage and one for outdoor purposes so I dont have to worry in case of smartphone theft
(31) Along with taking care of threats within the device such as phishing emails, I also take care of threats outside the device, for example device observations by surrounding people
(32) I get annoyed when I catch someone looking over my device screen
(33) I often catch people looking over my device screen without permission which irritates me
(34) I would like my device to do something everytime someone looks at it without my permission
(35) I hide my screen when I am in public areas
(36) I hide my screen when I am viewing sensitive information on my phone
(37) I scroll quickly when I sense someone is looking at my phone's screen
(38) I switch off my phone when I sense someone is looking at my phone's screen
(39) I don't view sensitive messages, play sensitive voice messages, or view perform sensitive actions (e.g., online banking) on my phone when I am in a public area
(40) I place my palm on touchscreens after i have entered sensitive information, to reduce the chances for thermal attacks to succeeed
(41) I press random keys on touchscreens to add noise to thermal imaging data
(42) I wipe my phone's screen with a cloth to prevent smudge attacks
(43) I don't leave my phone unattended to make sure no one attempts to use it or try to unlock it
(44) I keep my phone near me and visible to me all the time to make sure it is not compromised

A.3 Items Constructed Through Deductive Approach - Items by Larger Pool of Researchers

(45) I ensure no one is looking at my screen when I am entering passwords
(46) I use separate devices for private and non-private stuff
(47) I get anxious when someone from my surrounding invades my device privacy
(48) I am worried that someone might access my information by spying on what I am doing on my smartphone
(49) I am a privacy-centred person

(50) I would be embrassed if information found on my smartphone is leaked to my surrounding people
(51) I keep myself updated on how someone around me can unlock my smartphone without my permission
(52) I consider my data as a target from device external threats
(53) I believe my data is worth protecting from device external threats
(54) I protect my device from being observed by others
(55) I believe there are no data privacy threats outside of the device
(56) I am well aware of how to protect my data from device external threats
(57) I believe device external threats are not a serious privacy threat to be concerned of
(58) Device external threats are not effective in leaking private information
(59) I value protecting information from device external threats
(60) Device external threats are not a concerning threat to privacy
(61) It is fine for me to unveil my phone screen to the public
(62) It is important for me to protect from screen content from people around me in public transport
(63) I do not believe that someone could use my screen traces to attack my phone
(64) I dont mind if someone sitting next to me takes a look at my smartphone while I am watching a video
(65) I am concerned by the CCTC cameras as they might capture what I am doing on my device
(66) I protect my data from surrounding people
(67) I take all actions to keep my data safe from device external threats

B ITEMS USED IN THE PRE-TESTING PHASE

(1) I try to adjust my hand position when using my smartphone so that no one can see the information shown on it
(2) I hide my smartphone screen with my hands if I notice anyone looking at it
(3) If anyone looks at my screen without permission, I tend to put my smartphone away
(4) To avoid people nearby from looking at my smartphone screen, I specifically use a privacy-protecting screen cover (e.g. tampered glass protector)
(5) I lower my screen brightness so that no one around me can take a look at what is shown on it
(6) If someone sees my friend's content on my screen, it feels like a breach of my friend's trust in me to keep their content private
(7) I use fingerprint (or other biometric methods) mainly to avoid someone observing my password
(8) I trust the security system of smartphones and therefore store any sensitive information on them
(9) I would change all of my passwords immediately if my smartphone was lost
(10) I feel concerned when using ATMs that use cameras for recording purposes
(11) I wear gloves to avoid anyone taking a heat-trace picture of my PIN when I use an ATM
(12) I press extra keys after I have entered my PIN at the atm to avoid anyone taking a heat-trace picture of my PIN
(13) I access all sorts of data on my smartphone, including sensitive data
(14) Among the reasons I use auto-fill for passwords on my smartphone, is to avoid anyone overseeing what I enter
(15) I check for any surrounding people when I am doing something on my smartphone in public places
(16) The increasing availability and affordability of audio, video and photo recording devices are a threat to everyone's privacy
(17) I would like my device to do something to alert me every time someone looks at it without my permission
(18) I scroll quickly when I sense someone is looking at my smartphone's screen
(19) I place my palm on touchscreens after I have entered sensitive information, to reduce the chances for thermal attacks to succeed
(20) I press random keys on touchscreens to add irrelevant signals to thermal imaging data
(21) I ensure no one is looking at my screen when I am typing in passwords
(22) I consider my data to be a target for external threats to my device such as shoulder surfing
(23) I believe my data is worth protecting from external threats to my device
(24) Privacy invasion by people surrounding us is effective in leaking information
(25) It is not fine for me to have my smartphone screen visible to the public
(26) It is important for me to protect my screen content from people around me on public transport
(27) I believe that someone could use any finger tip traces on my screen to reveal my password
(28) I mind if a stranger sitting next to me takes a look at my smartphone while I am watching a private video
(29) I am a privacy-centred person
(30) I am worried that someone might access my information by spying on what I am doing on my smartphone
(31) I get anxious when someone from my surroundings invades my privacy by looking at the screen

C ITEMS EXPLORED IN THE EXPLORATORY FACTOR ANALYSIS

(1) I try to adjust my hand position when using my smartphone so that no one can see the information shown on it
(2) If anyone looks at my screen without permission, I tend to put my smartphone away
(3) To avoid people nearby from looking at my smartphone screen, I specifically use a privacy-protecting screen cover (e.g. tampered glass protector)
(4) I lower my screen brightness so that no one around me can take a look at what is shown on it
(5) If someone sees my friend's content on my screen, it feels like a breach of my friend's trust in me to keep their content private

(6) I use fingerprint (or other biometric methods) mainly to avoid someone observing my password
(7) I trust the security system of smartphones and therefore store any sensitive information on them
(8) I would change all of my passwords immediately if my smartphone was lost
(9) I feel concerned when using ATMs that use cameras for recording purposes
(10) I press extra keys after I have entered my PIN at the atm to avoid anyone taking a heat-trace picture of my PIN
(11) I access all sorts of data on my smartphone, including sensitive data
(12) Among the reasons I use auto-fill for passwords on my smartphone, is to avoid anyone overseeing what I enter
(13) I check for any surrounding people when I am doing something on my smartphone in public places
(14) The increasing availability and affordability of audio, video and photo recording devices are a threat to everyone's privacy
(15) I would like my device to do something to alert me every time someone looks at it without my permission
(16) I scroll quickly when I sense someone is looking at my smartphone's screen
(17) I consider my data to be a target for external threats to my device such as shoulder surfing
(18) I believe my data is worth protecting from external threats to my device
(19) Privacy invasion by people surrounding us is effective in leaking information
(20) It is not fine for me to have my smartphone screen visible to the public
(21) It is important for me to protect my screen content from people around me on public transport
(22) I believe that someone could use any finger tip traces on my screen to reveal my password
(23) I mind if a stranger sitting next to me takes a look at my smartphone while I am watching a private video
(24) I am a privacy-centred person
(25) I am worried that someone might access my information by spying on what I am doing on my smartphone
(26) I get anxious when someone from my surroundings invades my privacy by looking at the screen

D FINAL SET OF ITEMS & THE RESPECTIVE SOURCES

The table below presents the items from the final version of the out-of-device Privacy Scale and lists the corresponding sources from which the items were derived.

E EXPLORING MULTI-FACTOR SOLUTIONS - ADDITIONAL ANALYSIS

To finalize the factor solution, we explored factor solutions using direct oblimin (oblique) as the rotation method. We present and discuss the results below.

First, we explored a four-factor solution using 0.4 as the recommended loading cut-off value. The Table 5 below presents the results. It can be observed that no item is loaded onto the fourth factor. Therefore, we next explored a three-factor solution. Table 6 shows the output of a 3-factor solution. It can be observed that only two items are loaded onto the second factor, whereas at least three items must be loaded onto a factor for it to be considered a factor. Therefore, we dropped the three-factor solution and next explored a two-factor solution. The 2-factor solution (presented in Table 7 gave a simple structure; however, before finalizing it, we checked for the following descriptives:

(1) Correlation between the two factors: The correlation between the two factors turned out to be .553, indicating a high correlation.
(2) Reliability: We then checked for reliability, which appeared to be 0.857 for the first and 0.664 for the second factors. While the first factor gave a good reliability score, the reliability of the second factor was unacceptable.

While the above recommends opting for a single-factor solution, we further explored essential statistics. We collected a new dataset with N=1000 participants. Out of N=1000, 69 failed the attention check and were removed from further analysis. On the remaining N=931 participants' data, we performed the following tests. We again checked for a correlation between the two factors in the new dataset collected. The correlation between the two factors in the latest dataset was 0.590, indicating a high correlation. We then extracted loadings using Principle Axis Factoring (PAF) and CFA (Confirmatory Factor Analysis) for the two factors. For the loadings received using PAF (two-factor solution), the average variance extracted for each factor was 0.469 and 0.234, respectively. We then checked for the square root of AVE and compared it to the correlation. The square root of AVE was higher for only one factor (0.684) and not the other factor (0.483). For the loadings received using CFA, the average variance extracted for each factor was 0.47817 and 0.3207, respectively. We then checked for the square root of AVE and compared it to the correlation. The square root of AVE was higher for only one factor (0.691) and not the other factor (0.566). In both cases, insufficient discriminant validity was observed as factor correlation was not lower than the square root of AVE for Factor 2. Further, the AVE for each factor is less than 0.5, which is unacceptable as greater than 0.5 is the recommended threshold. Even with all these methods, TLI remains below the threshold of 0.9 (0.872). Given all these results, we opted for a single-factor solution.

Item Statement	Literature-Based	Reference	Empirically (by Researchers)	Empirically (by Experts)
It is important for me to protect my screen content from people around me on public transport				✓
I am worried that someone might access my information by spying on what I am doing on my smartphone				✓
I get anxious when someone from my surroundings invades my privacy by looking at the screen				✓
I am a privacy-centred person				✓
I consider my data to be a target for external threats to my device such as shoulder surfing				✓
I believe my data is worth protecting from external threats to my device				✓
It is not fine for me to have my smartphone screen visible to the public				✓
I mind if a stranger sitting next to me takes a look at my smartphone while I am watching a private video				✓
Privacy invasion by people surrounding us is effective in leaking information				✓
The increasing availability and affordability of audio, video and photo recording devices are a threat to everyone's privacy			✓	
I would like my device to do something to alert me every time someone looks at it without my permission			✓	
I would change all of my passwords immediately if my smartphone was lost	✓	[46]		
If someone sees my friend's content on my screen, it feels like a breach of my friend's trust in me to keep their content private	✓	[22]		
Among the reasons I use auto-fill for passwords on my smartphone, is to avoid anyone overseeing what I enter			✓	
I scroll quickly when I sense someone is looking at my smartphone's screen			✓	
I try to adjust my hand position when using my smartphone so that no one can see the information shown on it	✓	[23]		
I check for any surrounding people when I am doing something on my smartphone in public places			✓	
If anyone looks at my screen without permission, I tend to put my smartphone away	✓	[23]		

Table 4: The Table shows the list of items included in the final version of the ODPS and the corresponding sources.

Item Statements	**Factor**			
	1	**2**	**3**	**4**
I am worried that someone might access my information by spying on what I am doing on my smartphone	0.726			
I consider my data to be a target for external threats to my device such as shoulder surfing	0.701			
I am a privacy-centred person	0.546			
Privacy invasion by people surrounding us is effective in leaking information	0.49			
The increasing availability and affordability of audio, video and photo recording devices are a threat to everyone's privacy	0.455			
I believe my data is worth protecting from external threats to my device	0.447			
I mind if a stranger sitting next to me takes a look at my smartphone while I am watching a private video	0.443			
I feel concerned when using ATMs that use cameras for recording purposes				
It is not fine for me to have my smartphone screen visible to the public				
I would like my device to do something to alert me every time someone looks at it without my permission				
Among the reasons I use auto-fill for passwords on my smartphone, is to avoid anyone overseeing what I enter		0.517		
To avoid people nearby from looking at my smartphone screen, I specifically use a privacy-protecting screen cover (e.g. tampered glass protector)		0.448		
I use fingerprint (or other biometric methods) mainly to avoid someone observing my password		0.406		
I lower my screen brightness so that no one around me can take a look at what is shown on it				
I would change all of my passwords immediately if my smartphone was lost				
I check for any surrounding people when I am doing something on my smartphone in public places			-0.681	
If anyone looks at my screen without permission, I tend to put my smartphone away			-0.63	
I scroll quickly when I sense someone is looking at my smartphone's screen			-0.565	
I try to adjust my hand position when using my smartphone so that no one can see the information shown on it			-0.526	
I get anxious when someone from my surroundings invades my privacy by looking at the screen	0.432		-0.459	
It is important for me to protect my screen content from people around me on public transport			-0.445	
If someone sees my friend's content on my screen, it feels like a breach of my friend's trust in me to keep their content private				

Table 5: The table shows the results of a 4-factor solution using a loading cut-off value of 0.4.

Item Statements	Factor		
	1	2	3
I consider my data to be a target for external threats to my device such as shoulder surfing	0.651		
I believe my data is worth protecting from external threats to my device	0.57		
I am worried that someone might access my information by spying on what I am doing on my smartphone	0.553		
The increasing availability and affordability of audio, video and photo recording devices are a threat to everyone's privacy	0.523		
Privacy invasion by people surrounding us is effective in leaking information	0.492		
It is not fine for me to have my smartphone screen visible to the public	0.452		
I am a privacy-centred person	0.447		
It is important for me to protect my screen content from people around me on public transport	0.443		-0.41
I feel concerned when using ATMs that use cameras for recording purposes	0.423		
I mind if a stranger sitting next to me takes a look at my smartphone while I am watching a private video			
I would like my device to do something to alert me every time someone looks at it without my permission			
I would change all of my passwords immediately if my smartphone was lost			
Among the reasons I use auto-fill for passwords on my smartphone, is to avoid anyone overseeing what I enter		0.526	
To avoid people nearby from looking at my smartphone screen, I specifically use a privacy-protecting screen cover (e.g. tampered glass protector)		0.449	
I use fingerprint (or other biometric methods) mainly to avoid someone observing my password			
I lower my screen brightness so that no one around me can take a look at what is shown on it			
I check for any surrounding people when I am doing something on my smartphone in public places			-0.66
I scroll quickly when I sense someone is looking at my smartphone's screen			-0.617
I try to adjust my hand position when using my smartphone so that no one can see the information shown on it			-0.589
I get anxious when someone from my surroundings invades my privacy by looking at the screen			-0.555
If anyone looks at my screen without permission, I tend to put my smartphone away			-0.501
If someone sees my friend's content on my screen, it feels like a breach of my friend's trust in me to keep their content private			

Table 6: The Table shows the 3-factor solution using a loading cut-off value of 0.4.

	Factor	
Item Statements	**1**	**2**
It is important for me to protect my screen content from people around me on public transport	0.673	
I am worried that someone might access my information by spying on what I am doing on my smartphone	0.649	
I am a privacy-centred person	0.632	
I consider my data to be a target for external threats to my device such as shoulder surfing	0.63	
I believe my data is worth protecting from external threats to my device	0.619	
I get anxious when someone from my surroundings invades my privacy by looking at the screen	0.619	
The increasing availability and affordability of audio, video and photo recording devices are a threat to everyone's privacy	0.551	
I mind if a stranger sitting next to me takes a look at my smartphone while I am watching a private video	0.545	
It is not fine for me to have my smartphone screen visible to the public	0.539	
Privacy invasion by people surrounding us is effective in leaking information	0.533	
I feel concerned when using ATMs that use cameras for recording purposes		
I would like my device to do something to alert me every time someone looks at it without my permission		
If anyone looks at my screen without permission, I tend to put my smartphone away		
I would change all of my passwords immediately if my smartphone was lost		
If someone sees my friend's content on my screen, it feels like a breach of my friend's trust in me to keep their content private		
Among the reasons I use auto-fill for passwords on my smartphone, is to avoid anyone overseeing what I enter		0.557
I lower my screen brightness so that no one around me can take a look at what is shown on it		0.468
I scroll quickly when I sense someone is looking at my smartphone's screen		0.468
To avoid people nearby from looking at my smartphone screen, I specifically use a privacy-protecting screen cover (e.g. tampered glass protector)		0.465
I try to adjust my hand position when using my smartphone so that no one can see the information shown on it		0.444
I check for any surrounding people when I am doing something on my smartphone in public places		
I use fingerprint (or other biometric methods) mainly to avoid someone observing my password		

Table 7: The Table shows the 2-factor solution using a loading cut-off value of 0.4