# Self-Confidence Trumps Knowledge: A Cross-Cultural Study of Security Behavior

**Yukiko Sawaya** [*†] **Mahmood Sharif** [*‡] **Nicolas Christin** [‡] **Ayumu Kubota** [†]
**Akihiro Nakarai** [†] **Akira Yamada** [†]

## ABSTRACT

Computer security tools usually provide universal solutions without taking user characteristics (origin, income level, ...) into account. In this paper, we test the validity of using such universal security defenses, with a particular focus on culture. We apply the previously proposed Security Behavior Intentions Scale (SeBIS) to 3,500 participants from seven countries. We first translate the scale into seven languages while preserving its reliability and structure validity. We then build a regression model to study which factors affect participants' security behavior. We find that participants from different countries exhibit different behavior. For instance, participants from Asian countries, and especially Japan, tend to exhibit less secure behavior. Surprisingly to us, we also find that actual knowledge influences user behavior much less than user self-confidence in their computer security knowledge. Stated differently, what people think they know affects their security behavior more than what they do know.

### ACM Classification Keywords

H.1.2 User/Machine Systems: Human factors; K.4.4 Electronic Commerce: Security

### Author Keywords

Computer security; cross-cultural study

## INTRODUCTION

Understanding people's attitudes toward computer security is essential to devise effective human-centered defenses. We posit that cultural differences may considerably impact those attitudes, yet, the impact of these cultural differences has so far been under-studied in the realm of computer security. Most of the related studies indeed address either general cross-cultural considerations (e.g., [2, 3, 12, 32]), or very specific behaviors (e.g., smartphone locking [22]), thus motivating the questions: Does culture affect (general) computer-security behavior? If so, to which extent and how? We believe that answers to these questions are vital to shed light on the generalizability of previous work in computer security, and to guide future system design and interventions to help users be more secure.

In an effort to remedy that gap, this paper proposes to apply the previously proposed Security Behavior Intentions Scale (SeBIS, [16]) to participants from different countries, and to investigate the differences the scale could potentially reveal.

To do so, we first need to ensure that the SeBIS scale is properly translated, and that the concepts it attempts to measure carry over different languages and cultures. By attempting to translate the SeBIS scale into Japanese, and testing it with 1,654 users, we find that a mere translation, regardless of grammatical correctness, does not work: that is, statistical tests of model fit lead to rejecting the model previously validated for English speakers. Focusing on potential reasons for this discrepancy (and testing multiple hypotheses with additional participant pools), we discover that the original English scale uses double-negations and presents some ambiguities, which could have led to this negative result. We then devise a revised version of the SeBIS scale, and verify that the translation of the revised scale is robust across languages—that is, model fit tests are consistently satisfied. We publish the revised version of SeBIS in seven languages for the use of researchers and practitioners.

We subsequently use our revised scale in a survey of 3,500 participants from seven countries comprising about 29.15% of the world's population (500 participants in each of China, France, Japan, Russia, South Korea, the United States, and the United Arab Emirates). To the best of our knowledge, ours is the first study of user security-behavior on this scale. We build a regression model to study which factors affect participants' security behavior. Echoing and complementing earlier findings (e.g., [22]), we find that participants from different countries exhibit different behavior. For instance, participants from Asian countries, and especially Japan, tend to exhibit less secure behavior.

Surprisingly to us, we also discover that actual knowledge influences user behavior much *less* than users' self-confidence in their computer security knowledge. Stated differently, what people think they know affects their security behavior more than what they do know. Counter-intuitively, this indicates that a user with only passing security knowledge but high confidence may be more secure than a savvier but less confident user. We believe that this (re)opens avenues

[*]The first two authors contributed equally to this work.
[†]KDDI Research, Inc., Saitama, Japan.
{yu-sawaya, kubota, ak-nakarai, ai-yamada}@kddi-research.jp
[‡]Carnegie Mellon University, Pittsburgh, PA, USA.
{mahmoods, nicolasc}@cmu.edu


*CHI 2017, May 6-11, 2017, Denver, CO, USA.*
ACM ISBN 978-1-4503-4655-9/17/05.
http://dx.doi.org/10.1145/3025453.3025926

(e.g., individualized messaging) to help users improve their security.

Next, we present related work. Then, we discuss the development process of the revised SeBIS scale. Subsequently, we present our main experiment, studying the the effect of culture, among other factors, on computer security behavior. We end with a discussion and concluding remarks.

## RELATED WORK

A number of previous efforts are directly related to our work. We first present work on the development of cross-cultural surveys, which will inform the development of our own survey. We then discuss the results of cross-cultural surveys that are related to the area of computer security and privacy. Then, we present the main results of studies that explored the relationship between security knowledge and behavior. Finally, we present the Security Behavior Intentions Scale and the results of efforts made to test its ability to predict actual behavior.

### Cross-Cultural Surveys

Although translation may seem as a straightforward task, not taking appropriate precautions when translating a scale may result in invalid measurements (e.g., [13, 21]). For example, some words and concepts (such as the concept of *dépaysement* in French, related to the sense of disorientation one may feel in a foreign or new environment) exist in some languages but not others, and could be misinterpreted if not translated carefully. A failure in preserving the meaning of questions after translation may cause the answers of participants from different countries to be incomparable. Due to the ubiquity and pervasiveness of computers and mobile devices, concepts from computer security and privacy are widely understood. However, not paying attention to translation subtleties may still lead to failure in maintaining the validity and reliability of the scale [45].

In particular, the effect of reverse-worded questions (questions whose answers are inversely correlated with the trait being measured by the scale) on the reliability of scales when applied across cultures has been studied widely [8, 51, 55]. Such questions are often used in scales to prevent response bias caused by participants who follow a specific response style rather than faithfully answering questions [51]. In the context of cross-cultural studies, however, it is widely recommended to avoid reverse-worded questions as they may lead to unexpected factor structures (i.e., harming the construct validity of scales), especially due to the possibility of being misinterpreted by participants [8, 55]. As we show later, rephrasing reverse-worded questions is integral to our reliably translating the scales we use from English to other languages.

Other work suggests to validate the correctness of translations by translating questions from the target languages back to the original language and ensuring that they are semantically identical to the original version [22]. For example, to validate that the translation of a question from English to Russian is correct, one can ask a Russian speaker to translate the question back to English so that it can be compared with the original version. We employ this technique to comprehensively validate the correctness of our translations.

### Culture in Computer Security and Privacy

Cross-cultural studies in the area of computer security and privacy studied the differences in privacy concerns and behavior between different cultures and explored the reasons underlying these differences (e.g., [2, 3, 12, 32, 53, 57]). For example, Almakrami compared the self-disclosure practices of Saudis on Facebook with those of Australians [2]. Through a quantitative study, he found that whereas Australians tend to be conservative with their online disclosure relatively to their openness in offline relationships, Saudis tend to be more open and free on Facebook than in their offline relationships. Subsequently, Almakrami ran a qualitative study to discover the cause behind these differences and concluded that they may result from the restrictions on offline relationships that exist in Saudi Arabia, but not in Australia.

In the late 60s and early 70s, Geert Hofstede surveyed more than 110,000 employees of IBM across 40 countries and identified four dimensions of cultural values: power distance, individualism, uncertainty avoidance, and masculinity [23]. In a later stage, he added two more dimensions (long-term orientation and indulgence) [24]. A large body of work has used the cultural-value dimensions and showed that they correlate with various cultural traits. In fact, some studies explored how Hofstede's dimensions of cultural values correlate with privacy concerns [4, 7]. However, other work criticizes Hofstede's theory and methodology for being inconsistent [1] and over generalizing [37]. Therefore, we decided not to use Hofstede's dimensions of cultural values, and rather use nationality as a proxy for culture.

Various research efforts compared the privacy (but not security) concerns of Internet users from Asian countries with those of their counterparts from western countries [27, 29, 34, 35, 53, 52, 57]. In general, most of these efforts concluded that the Asian users are less concerned about their online privacy than users from western countries. Nevertheless, many found that Asian users' lower levels of concern do not necessarily lead them to disclose more information online. For instance, Wang et al. ran an online study with 924 participants to compare the privacy concerns of American, Chinese, and Indian users on online social networks [52]. They found that American users were more concerned than Chinese users, while the Indian users were the least concerned. At the same time, American users expressed the least desire to limit the visibility of their information from certain sets of people.

As a result of these cross-cultural privacy studies, researchers suggested that privacy policies and tools, rather than being static, should adapt themselves to their users while taking culture into account [32, 50]. Ur and Wang proposed a "check list" of questions that online social networks should address to ensure that they provide reasonable privacy protections to users from diverse cultural background [50]. One question that they propose online social networks should address is whether they are aware of the cases in which data revelation may cause distress to users from certain cultures. Krasnova et al. suggested that online social networks should explicitly

advise international users about the origin and implications of the legal framework with which they comply [32].

In the realm of computer security, a few studies compared perceptions and behavior of people from different cultures [5, 22, 30]. Karvonen et al. ran an interview study to compare Swedes' and Finns' perceptions of computer security e-commerce websites [30]. They found no particular differences, possibly due to the close geographical distance and the shared history between Sweden and Finland. In a study with 36 university students, Chaudhary et al. tested if culture influences understanding and self-assessment of awareness to security risks, with a particular focus on phishing [5]. They found that Finnish students overestimate their knowledge compared to Chinese students. Finally, Harbach et al. studied the differences in smartphone locking behavior among people from eight countries [22]. They discovered that users' perceptions of the sensitivity of their data and their security behavior contradict sometimes. For instance, the Japanese participants in their study considered the sensitivity of their smartphones' contents to be higher than participants from other countries. Yet, the Japanese participants were among the least likely to use a secure locking mechanism. Differently from these studies, our work does not focus on a specific aspect of computer security, but rather considers security behavior broadly. Moreover, we control for factors that were not considered by previous studies, such as the extent of knowledge on computer security that the participants have, and how confident they are in their knowledge.

### Knowledge and Security/Privacy Behavior

Although it seems natural to expect users who know more about computer security and those who are engaged in maintaining their devices' security to follow more secure behavior, previous work has shown that this may not always be the case [17, 26, 54]. Wash and Rader suggested that security knowledge does not always capture the range of security beliefs users have, and thus may fail in predicting their actual behavior [54]. Ion et al. observed that security experts may occasionally behave in ways that contradict with security advice that they often give to non-experts [26]. For instance, even though experts often recommend not opening emails from unknown people, they reported doing so more often than non-experts. Finally, Forget et al. found that the computers that belonged to users who are engaged in maintaining their security were often in less secure states than the computers of less engaged users [17].

Kraus et al. studied the extent to which concern and knowledge about security and privacy affect people's mobile protection behavior [33]. In contrast to the above mentioned work, they found that more concern and knowledge was correlated with higher usage of protection methods. A possible explanation for the difference is that their work focused on the protection of mobile phones whereas the other work studied security behavior in a broader sense. Kraus et al. also developed a test which consists of 11 multiple answer questions to evaluate knowledge of security and privacy concepts [33]. We consciously decided not to use their test because of its privacy focus (which is not the aim of our study) and due to possible response fatigue that the multiple answer questions may cause to participants. Instead, we use true/false questions to evaluate our participants' security knowledge.

### The Security Behavior Intentions Scale (SeBIS)

Egelman and Peer proposed the Security Behavior Intentions Scale (SeBIS) as a tool to facilitate the measurement end-users' security behavior [16]. Following a rigorous development method, they eventually converged to a scale that consists of 16 Likert-scale questions. Each question measures one of four security related factors: proactive awareness, password selection, device securement, and device updating.

We choose to use SeBIS in our study as it was shown not only to capture users' reported security behavior, but their *actual* behavior as well [14, 15]. In one work, Egelman and Peer found that SeBIS's device-updating sub-scale is significantly correlated with installing anti-virus or firewall software (users with higher score on the sub-scale are more likely to install such software). In a more recent work, they reported that also the other three sub-scales correlate with security behavior with users who score high on the "proactive awareness" sub-scale being more likely to correctly identify phishing websites, users who score high on "password selection" creating passwords that are harder to crack, and those who score highly on "device securement" being more likely to use a secure locking mechanism on their smartphones [14]. Since SeBIS has been shown to be highly correlated with actual security behavior, for simplicity, we often mention that we use SeBIS to estimate security behavior, when we actually mean security behavior-intentions.

## PRELIMINARY EXPERIMENT: SURVEY TRANSLATION

We next present our methodology for translating SeBIS. We follow this methodology to achieve a reliable and valid version of the scale in Japanese. We will later show that using the same techniques results in reliable and valid versions of the scale in other languages as well.

#### *Naive Translation*

We started by having the original, English, SeBIS scale translated into Japanese. The decision to start from Japanese was informed by the difficulty of translating English to Japanese [18], and the makeup of the research team, which made security experts who are proficient in both English and Japanese readily available to us. The goal of the translation was to preserve the original meaning of the questions to the best extent possible.

Relying on a survey company (Macromill, [36]), we then randomly selected a pool of 30,000 potential participants in Japan. We recruited 1,654 participants from this pool, so that our recruited sample's demographics (gender and age) statistically matches those of Japanese Internet users as reported in the 2010 national census [46]. The survey company administered the survey online on our behalf.

Similarly to the original SeBIS paper [16], we then use confirmatory factor analysis (CFA) and Cronbach's $\alpha$ to evaluate the validity and reliability of the translated SeBIS scale for Japanese users. Confirmatory factor analysis

measures the goodness of fit between the scales' items and a set of hypothesized latent factors (in this case: proactive awareness, password selection, device securement, and device updating). High goodness of fit shows the items measure the factors we are expecting them to measure, i.e., the scale is valid. Cronbach's alpha measures the scale's reliability, i.e., that the items are measuring the same construct. This is important, as an unreliable scale cannot be valid [47].

Table 1 summarizes our results. Let us focus for now on the two leftmost columns. The first column presents Cronbach's $\alpha$ reliability measure and several data fit indices (RMSEA, SRMR, CFI and TLI) for the original, English, SeBIS scale [16]. The second column presents the same indices for the literal Japanese translation ("preliminary experiment") of the SeBIS scale. Contrary to the English scale, the RMSEA and SRMR are both above the cutoff points recommended by Hu and Bentler [25]; and the CFI and TLI are both below the 0.90 cutoff recommended by Netemeyer et al. [40]. Thus, these statistical tests indicate poor model fit, and lead us to reject the literal translation of the SeBIS scale.

### *Possible Causes of Translation Failure*

We next investigate the causes for the poor model fit we observed. To do so, we formulate the following hypotheses:

- $H_1$: **Demographic differences between the two populations cause the poorness of fit.** The original SeBIS paper relied on participants aged 19 to 71 (average: 34.3, SD: 10.78) and 46.8% were female and 52.8% were male. On the other hand, our participant pool – aiming to be representative of the Internet demographics in Japan – ranges from 15 to 69 (average: 44.3, SD: 14.80, male: 51.5%, female: 48.5%). That is, our sample is slightly older on average than the population sample used by Egelman and Peer; if age negatively impacts security behavior, we would expect to get worse goodness of fit values.
- $H_2$: **Inadequate filtering of participants resulted in poor goodness of fit.** Contrary to Egelman and Peer [16], our initial translation did not feature attention questions aimed at discarding participants who answer without paying enough attention to their answer. We hypothesize that such "bogus" answers might have caused the poorness of fit.
- $H_3$: **Linguistic particularities impact the SeBIS scale.** Literal translation does not account for certain linguistic particularities. For instance, answers to negative interrogative sentences may differ between languages. Likewise, certain concepts (e.g., "security") may be more ambiguous in a language than in another. We hypothesize that the poorness of model fit is (at least partly) due to such particularities.

### *Hypothesis Testing*

To test $H_1$, we sampled a set of 505 participants from our larger participant pool, designed to match the demographics of the population sample Egelman and Peer used [16], that is, ages ranging from 18 to 69 (average: 34.6, SD: 10.87) and 44.8% were female and 55.2% were male. The results, shown in Table 1 (column $H_1$), indicate a poor fit, similar to our preliminary experiment. We reject $H_1$.

To test $H_2$, we recruited another cohort of 1,654 participants without overlap with those who participated in our preliminary experiment or in our test of $H_1$. We reran the test of the literal Japanese translation of the SeBIS scale, adding attention questions identical to those used by Egelman and Peer [16] to weed out participants who did not read questions carefully.

The results, presented in Table 1, indicate that the model fit remains poor, despite the inclusion of these attention questions. We thus rule out $H_2$.

We are left with $H_3$. We first performed factor analysis to determine if certain questions were more problematic than others. We discovered that reverse-worded items did not load on the same sub-scales they were assigned to by Egelman and Peer [16]. In short, reverse-worded questions appear to have been a possible cause of problems in our original translation.

One of the root causes may be the use of negative interrogation sentences, whose answers are opposite in English and Japanese. For example, to the question "aren't you tired?" an English speaker who does not feel tired would answer "no, I am not tired." A Japanese speaker, to denote the exact same state, would answer (literally translating) "yes, I am not tired." As a concrete example from SeBIS, "I do not change my passwords, unless I have to" is a problematic reverse worded item that we identified. Speakers of some languages (particularly, Japanese) may respond to it by either "yes, I do not" (thus answering "always" on SeBIS) or "no, I do not" (i.e., "never"). Rephrasing as "I change my passwords even if it is not needed" eschews this issue. Another possible issue we identified is that some of the reverse-worded questions appear more ambiguous when translated literally in Japanese, than they are in the original English phrasing. For example, the question "I know what website I'm visiting based on its look and feel, rather than by looking at the URL bar" combines somewhat abstract concepts ("based on its look and feel") with a more concrete proposition ("looking at the URL bar"). We conjecture that this shift may have confused Japanese speakers.

To test $H_3$, we rephrased all of the scale items that were originally reverse-worded so that they matched the logical order of the other items, and attempted to make them as unambiguous as possible. We call the revised scale Refined SeBIS (RSeBIS; see the Appendix for the English version, and the auxiliary material for versions in other languages).

We recruited another 1,654 participants, following the same demographics as in our preliminary experiment, and tested RSeBIS. The results, shown in Table 1 (column $H_3$) indicate high reliability and good fit, pretty much equivalent to the reliability and goodness of fit of the original scale in English. Since RSeBIS consists of four sub-scales (i.e., it is a multidimensional scale), we also measured $\alpha$ for each sub-scale to ensure they are reliable, following the suggestion of Tavakol and Dennick [47] and Egelman and Peer [16]. We found that $\alpha$ lies in the recommended range for each. Our results, thus, align with previous findings on how reverse-worded questions often degrade the reliability and validity of scales when applying them across cultures [8, 51, 55].

| Experiment | Original | Preliminary | $H_1$ | $H_2$ | $H_3$ | $H_4$ | |
|---|---|---|---|---|---|---|---|
| Scale | SeBIS(en) [16] | SeBIS(ja) | SeBIS(ja) | SeBIS(ja) | RSeBIS(ja) | RSeBIS(en) | Recommended |
| N | 500 | 1,654 | 505 | 1,654 | 1,654 | 408 | |
| Cronbach's $\alpha$ | 0.801* | 0.794* | 0.814* | 0.685* | 0.893* | 0.86* | >0.60 [16] |
| RMSEA | 0.058* | 0.087 | 0.089 | 0.095 | 0.062 | 0.061 | <0.06 [25] |
| SRMR | 0.050* | 0.090 | 0.097 | 0.107 | 0.042* | 0.054* | <0.08 [25] |
| CFI | 0.920* | 0.818 | 0.828 | 0.753 | 0.949* | 0.934* | >0.90 [40] |
| TLI | 0.902* | 0.777 | 0.789 | 0.698 | 0.938* | 0.919* | >0.90 [40] |

**Table 1. Reliability and model fit-indicators. The first column represents the values indicated by Egelman and Peer. The second to fourth column show values for tests across various population samples using a literal Japanese translation of the scale. The fifth column shows the values for our proposed revised scale in Japanese, and the sixth column shows the fit for our revised scale in English. Stars indicate good model fit.**

*Running RSeBIS in the US*

After retaining $H_3$, we wanted to validate that the refined scale remains effective in English. To this end, we rephrased reverse-worded questions in the original SeBIS scale, and ran the refined survey on Amazon's Mechanical Turk to test the following hypothesis.

**$H_4$: RSeBIS presents good model fit in English and for US populations as well.**

We recruited 408 participants from the US, with demographics comparable to those used in the original work [16]. We found that the reliability measures exceed the desired minimum expected from a reliable scale ($\alpha > 0.7$ for the entire scale, $\alpha > 0.6$ for each sub-scale). Further, as shown in Table 1, the values of all the goodness-of-fit measures are very close to, or well within the ranges of recommended values. Therefore, we retain $H_4$.

## MAIN EXPERIMENT: METHODOLOGY

Having identified and rectified potential issues with the SeBIS scale, we can turn to our main experiment. To study the effect of culture on computer security behavior, we followed a between-subject design. We developed an online survey in which we elicited information about security behavior and various other factors that may potentially affect it from participants located in seven countries. The survey was developed in English and then translated into six other languages. Below, we provide the details about the survey, the translation process, the sample we study, and we comment on the limitations of our methodology.

### Measuring Security Behavior

Users make many security-related decisions on a daily basis. These range from selecting passwords for their online and offline accounts, or locking their devices when they step away from them, to updating the anti-virus software they use, or deciding whether an email they have just received is a phishing email. While observing and measuring users' actual behavior would be ideal for the purpose of studying how various factors affect it, doing so at a scale, especially in a cross-cultural setting, is prohibitive.[1] In a previous attempt, for example, researchers were able to observe the security behavior of 73 users, all located in the same city [17]. Such sample is inadequate to address our research questions.

[1]Large companies with access to telemetry data, such as Symantec, may be an exception, but telemetry data alone does not allow to differentiate between "operator error" and actual security posture.

Here, we decided to use RSeBIS to measure users' security behavior. As explained in the previous section, RSeBIS is an equivalent of SeBIS [16] in which reverse-worded questions are rephrased to match the logical order to other questions. While users' responses to RSeBIS and SeBIS only indicate the frequency in which they engage in secure behavior, previous work has shown that the scale they measure is strongly correlated with users' actual behavior [14, 15]. Thus, after translating the scale, RSeBIS allows us to estimate actual behavior at scale and in different cultures.

### Measuring Factors Affecting Security Behavior

Culture may not be the only factor that affects security behavior. Other differences (such as age, education level, knowledge about security, ...) may also lead to differences in user behavior. Therefore, we also control for other factors that may explain behavioral differences. These factors include: knowledge about computer security, self-confidence in one's ability to keep his/her data and devices safe, and demographics.

*Nationality as a Proxy for Culture*

Culture is a complex term that is hard to describe, let alone quantify. Previous studies (e.g., [4, 7]) used indices of cultural values (such as Hofstede's [23]) to approximate it. These indices are often criticized to be a coarse over generalization [37]. To avoid possible confounds, we decided not to use indices of cultural-values. Instead, we follow the steps of previous work and use nationality as a proxy for culture [50, 52].

In this work, we study the security behavior of participants from seven countries: China, France, Japan, Korea, Russia, the United Arab Emirates (UAE), and the United States (US). This set of countries covers five different geographical regions (North America, East Asia, the Middle East, Eastern Europe, Western Europe), and consists of about 29.15% of the total world population. Moreover, each country in this set has a different official language. While selecting a relatively small set of languages always incurs the risk of inadvertently biasing the selection toward certain characteristics, we believe that our selection represents a relatively sound sample given the constraints we face. For instance, we ruled out India due to the multiplicity of different local languages, and the fact that an overwhelming majority of computer-literate people in India speak English.

*Testing Security Knowledge*

We developed a set of 18 true/false questions to test our participants' knowledge about computer security and secure

behavior. To develop this set of questions, we first compiled a list of 81 questions (in English) by rephrasing security advice that is given to users by Internet service providers and governmental bodies as questions, and by relying on our expertise in the area of computer security. Subsequently, we discarded questions that (a) were almost identical to other questions; (b) required deep technical knowledge (e.g., a question regarding the SMTP protocol); or (c) questions to which we did not have a complete agreement within ourselves for whether the correct answer is "true" or "false." The output of this filtering process is a set of 18 questions. We then rephrased some of the questions to eliminate negations, as translating negations could potentially lead to confusion (as described earlier). We refer the reader to the Appendix for the final set of questions in English. Versions in other languages can be found in the auxiliary material.

*Self-Confidence in Security Knowledge*

We estimate self-confidence in security knowledge using 5-point Likert-scale questions (from "strongly disagree" to "strongly agree"). In each question, participants report the extent to which they agree with statements regarding their knowledge about countermeasures that can be taken to prevent their devices, their data, or their money, from being compromised. We report the English version of the questions in the Appendix (See auxiliary material for versions in other languages).

*Demographics*

To control for demographic differences that may affect security behavior, we also collect information about our participants' demographics. The demographic information we collect includes: age, gender, whether the participant has a degree or a job in a technical area, the area in which the participant lives or grew up (urban, suburban, rural), and their income level.

Income levels are very disparate across the countries we consider. To remedy this, we rely on the "Atlas method," proposed by the World Bank [56]. That is, we normalize the income levels presented to participants by multiplying them with the ratio between the gross national income (GNI) per capita in their country and the GNI per capita in the United States. For example, a US income level of "$20K/year–$40K/year" is first converted to "$2.84K/year–$5.69K/year" for Chinese participants. This is done by multiplying the original US income level by the ratio between the "Atlas methodology values" [41] for China ($7.82K) and the United States ($54.96K). We then convert the income levels resulting from this normalization to amounts expressed in the official currency used in the participants' countries, based on the average currency rate in 2015; in our example, we obtain a range "CNY 17,800–35,500" once expressed in Chinese Yuan.

### Survey Translation

To translate the survey, we enlisted the help of a number of computer security experts from our team and among our acquaintances that are also native speakers of Arabic, French, Japanese, Chinese,[2] Korean, and Russian.

[2]We used simplified characters for the Chinese translation.

As a first step to validate our translations, we followed the approach suggested by Harbach et al. [22]. For each question, we compared five "reverse" translations (i.e., translating the translations back to English) with the original English version and validated that both are semantically equivalent. For all languages except Arabic, we crowdsourced reverse translations using Gengo [20], an online service for crowdsourcing translations. Unfortunately, Gengo does not offer translations from Arabic to English. Therefore, the Arabic translation did not pass through this preliminary validation step. With the exception of question #13 of the knowledge test, at least four out of five reverse translations of each question (from each one of the five languages) agreed with the original English version. Thus, we removed the 13th question of the knowledge test from the survey, and kept the remaining questions whose translation preserved the original meaning.

### Participants

We used the services of Cross Marketing Inc. [11] to collect responses to our survey. Cross Marketing Inc. is a Japanese company that offers online research services with a reach to a pool of 30 million participants located around the world. In total, we collected 3,500 responses to our survey; from each one of the seven countries studied, we collected 500 responses.

The number of participants was determined after running power analysis on the data collected in the preliminary experiment. Our analysis showed that, in the case of two countries, about 207 participants from each country are necessary to detect differences in security behavior intentions with probability of at least 0.8, when the chance of a Type-I error (i.e., significance level) is below 0.05. Since we aimed to compare participants from a larger number of countries, we sampled more than double this number from each country, as a safety margin.

Since our goal is to study computer security behavior, we aimed to collect samples of participants that are representative of technology consumers in the studied countries. Therefore, we instructed the survey company to match (to the best extent possible) the samples' age and gender to those of the populations of Internet users in the seven countries. Table 2 presents stats on the distribution of gender and age for each one of the countries. With the exception of the Chinese sample (which has a higher ratio of females and older participants than expected), our samples' age and gender distributions are relatively similar to those of the actual populations in their respective countries.

## MAIN EXPERIMENT: RESULTS

In this section, we present our results based on the analysis of the 3,500 responses that we collected. First, we provide results showing that the translations preserved the validity and reliability of RSeBIS. Then, we present our multiple regression model that sheds light on how multiple factors, and most importantly culture, affect security behavior.

### Evaluating the Translations

In the previous section, we showed that our translations of the survey were semantically equivalent to the original English

| | Gender | | | Age | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Country | Male% | Female% | Actual Male% | Average | Median | SD | Min | Max | Actual Median |
| China | 48% | 52% | 56% [6] | 42.44 | 39 | 13.09 | 19 | 89 | 20–29[6] |
| France | 49% | 51% | 49% [10] | 41.81 | 40 | 15.43 | 18 | 82 | 40–59[10] |
| Japan | 54% | 46% | 50% [9] | 36.83 | 34 | 13.92 | 18 | 77 | 35–44[9] |
| Korea | 51% | 49% | 52% [31] | 43.49 | 43 | 13.54 | 19 | 78 | 35–44[31] |
| Russia | 50% | 50% | 47% [49] | 43.29 | 44 | 11.47 | 18 | 76 | 35–44[49] |
| UAE | 65% | 35% | 72% [28] | 32.89 | 32 | 8.37 | 18 | 60 | 31–40[28] |
| USA | 49% | 51% | 49% [42] | 43.83 | 42 | 17.84 | 18 | 91 | 30–39[42] |

**Table 2. The age and gender of our each of our samples. The fourth column (from the left) shows the actual percentage of men within the population of Internet users. The last column shows the range in which the actual median age of Internet users lies.**

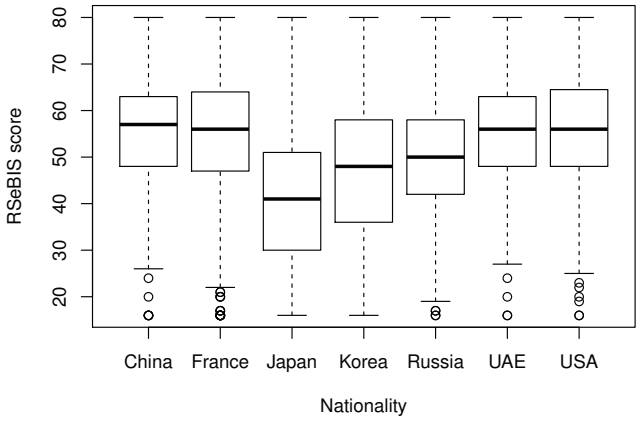


**Figure 1. A box plot showing the RSeBIS scores of participants in each country studied. The middle line in each box is the median. The bottom and top of each box extend to the 25th and 75th percentile, respectively. The whiskers of boxes extend up to $\times 1.5$ the interquartile range (points outside this range are considered outliers).**

version. However, our preliminary experiment with the Japanese translation, which motivated our design of RSeBIS in the first place, had shown that semantic equivalence did not necessarily result in a reliable scale. So, we tested the reliability and the model fit of RSeBIS for each one of the seven countries. Similarly to Egelman and Peer [16], and to our preliminary experiment, we used Cronbach's $\alpha$ indicator to measure reliability, and the RMSEA, SRMR, CFI and TLI indicators to measure model-fit.

Table 3 summarizes the results. In all languages, RSeBIS has high reliability (for the entire entire scale, and each of its sub-scales), and at least two out of the four model-fit indicators show its model fit lies in the recommended range. Thus, these results provide another confirmation that our translations are reliable and effective.

### Factors Affecting Security Behavior

After establishing basic trust in the quality of our translation, we turned our focus to the main research question: how does culture (among other factors) affect security behavior?

By exploring the RSeBIS scores of participants from different countries, one can see some trends (as shown in Figure 1). Each RSeBIS score is a natural number in the range between 16 to 80 where a higher score indicates more secure behavior. We can notice that the RSeBIS scores of Japanese participants are remarkably lower than the scores of their counterparts. On the other hand, American, Chinese, Emirati, and French participants seem to exhibit more secure behavior than other participants.

However, a simplistic analysis of how culture affects security behavior independently from other explanatory variables may not lead to an accurate characterization of the differences in security behavior between cultures. To this end, we use a multiple (linear) regression model, that enables us to study the effect of culture on security behavior while controlling for other factors. Specifically, we use the following model:

$$\texttt{RSeBIS} \backsim \texttt{Gender} + \texttt{Income level} + \texttt{Area} + \texttt{Culture} + \texttt{Knowledge}$$

`Gender`, `Income level`, and `Area` are explanatory variables related to demographics. `Gender` is an indicator variable to indicate if the participant is a male or not (female is the baseline). `Income level` consists of several indicator variables for different range of income. The baseline range is \$0K per year to \$20K per year (before correction by the Atlas method value for each country, as detailed in the methodology). Each subsequent level then increases by (the equivalent of) \$20K. We note that 10% of the participants preferred not to provide their income level. Our model includes an indicator variables to account for this preference. `Area` indicates the type of area in which the participant lives or grew (the baseline is Urban). 98 participants (2.80% of the total) who decided not to provide this information were dropped when estimating the coefficients of the model. We did not include age in the model as our initial analysis showed that it does not correlate with RSeBIS (even when splitting the data by country). In addition, including it in the model lead to almost no difference in the results.

`Culture` consists of indicator variables for the nationality of the participants. We set the USA as the baseline level as its population is studied more often than other nations' populations. This choice aims to facilitate the comparison of our results with related work.

`Knowledge` consists of three different variables. The first variable is the score in the knowledge test, which is a natural number in the range 0 to 17. The second variable is the self confidence in knowledge. This is a natural number in the range 5 to 30. Finally, the third variable is an indicator variable for

| | China | France | Japan | Korea | Russia | UAE | USA | Recommended |
|---|---|---|---|---|---|---|---|---|
| Cronbach's alpha | 0.895* | 0.896* | 0.928* | 0.906* | 0.855* | 0.852* | 0.881* | >0.600 [16] |
| RMSEA | 0.069 | 0.068 | 0.064 | 0.072 | 0.066 | 0.061 | 0.057* | <0.060 [25] |
| SRMR | 0.058* | 0.051* | 0.041* | 0.044* | 0.061* | 0.044* | 0.044* | <0.080 [25] |
| CFI | 0.928* | 0.937* | 0.960* | 0.937* | 0.925* | 0.925* | 0.947* | >0.900 [40] |
| TLI | 0.903* | 0.912* | 0.943* | 0.916* | 0.896 | 0.891 | 0.918* | >0.900 [40] |

**Table 3. Reliability and model-fit indicators for RSeBIS. Stars indicate values that lie in the recommended ranges.**

| Indep. Vars | Estimate | *p*-values | | 95% CI |
|---|---|---|---|---|
| (Intercept) | 26.610 | <0.001 | *** | [24.144, 29.076] |
| Male | -0.237 | 0.543 | | [-1.001, 0.527] |
| Income$20–40K | 0.628 | 0.284 | | [-0.522, 1.778] |
| Income$40–60K | 1.119 | 0.067 | | [-0.080, 2.318] |
| Income$60–80K | 3.401 | <0.001 | *** | [2.053, 4.749] |
| Income$80–100K | 3.562 | <0.001 | *** | [1.945, 5.179] |
| Income$100K+ | 3.731 | <0.001 | *** | [2.336, 5.126] |
| IncomeNA | 0.132 | 0.850 | | [-1.233, 1.497] |
| GrewInSuburb | 1.018 | 0.117 | | [-0.256, 2.292] |
| GrewInRural | -0.135 | 0.810 | | [-1.232, 0.963] |
| LivesInSuburb | 0.167 | 0.789 | | [-1.055, 1.390] |
| LivesInRural | 0.209 | 0.765 | | [-1.161, 1.578] |
| KnowTestScore | 0.265 | <0.001 | *** | [0.136, 0.395] |
| KnowSelfConf | 1.138 | <0.001 | *** | [1.065, 1.210] |
| Technical | 3.127 | <0.001 | *** | [2.211, 4.042] |
| Chinese | -1.909 | 0.012 | * | [-3.395, -0.423] |
| Emirati | 0.571 | 0.470 | | [-0.978, 2.120] |
| French | 1.717 | 0.017 | * | [0.306, 3.127] |
| Japanese | -9.199 | <0.001 | *** | [-10.636, -7.762] |
| Korean | -4.548 | <0.001 | *** | [-6.011, -3.084] |
| Russian | -4.148 | <0.001 | *** | [-5.626, -2.671] |

**Table 4. Coefficient estimates of the linear regression model. *p*-values test the hypothesis that coefficients are zero (i.e., exploratory variables not affecting RSeBIS). Significance codes: '***': $p<0.001$, '*': $p<0.05$. The rightmost column presents the 95% confidence intervals. Actual income levels are mapped to the corresponding US incomes by the World Bank's Atlas method.**

whether the participant works or holds a degree in a technical area.

Table 4 presents the coefficient estimates of our model. We note that, before estimating the coefficients, we validated that the linear regression's assumptions (linearity, fixed-$x$, and independent errors) hold. Post-analysis of the residuals showed no violation of the assumptions regarding linearity or error's normal distribution. Moreover, correlation analysis (using Spearman's rank correlation test) between our explanatory variables showed no particular correlation between pairs of variables. Therefore, we believe that our results are not affected by collinearity. Our estimate's coefficient of determination (also called $R^2$) is 0.385. In other words, our model explains 38.50% of the variance in the data. This value is considered to be in an acceptable range (despite different perspectives on the usefulness of $R^2$ [38]). We now explain what we learn from our model.

### *Culture*

By examining the linear regression's estimated coefficients, we can learn that there are stark differences in security behavior between participants from different cultures. For instance, French participants exhibited the most secure behavior. The estimated mean RSeBIS score of our French participants is higher by 1.717 than that of our American participants. The estimated means RSeBIS score of our Emirati participants and of our American participants are not significantly different from each other. On the other hand, our Russian participants produced a mean RSeBIS score lower by 4.148 than that of their American counterparts.

Of particular note is that the estimated mean RSeBIS scores of participants based in the Asian countries we surveyed are significantly lower than the baseline. Specifically, the Japanese participants exhibited remarkably less secure behavior than other participants. Their estimated mean RSeBIS score is lower by about 9 than that of the American participants. Our results align with previous work that shows that Asian users are likely to be less concerned about online privacy, to exhibit less private behavior online, and not to use phone locking-mechanisms, than their western counterparts (e.g., [22, 29, 35, 53]).

To put our results in perspective, one can contrast them with Egelman et al.'s mapping between SeBIS scores and actual behavior [14]. They found that users with strong passwords score an average of 1.04 points on the password sub-scale higher than users with weaker passwords; users who can detect phishing websites score an average of 3.15 points higher on the awareness sub-scale than users who cannot; and users who regularly update their systems score an average of 1.5 points higher on the updating sub-scale than users who do not. Altogether, this adds up to 5.69, less than the mean RSeBIS score difference between several populations in our study. For example, Japanese and American participants only differing in nationality are roughly 9 points apart. Hence, we can conclude that users from some cultures (especially, users from Asian cultures) are more likely to exhibit insecure behavior than users from other cultures.

### *Demographics*

Our model does not shed light on any particular differences in security behavior between men and women. Similarly, the effect of the area in which people grow up or live in seems to have almost no effect effect on how securely they behave. In contrast, income has a large effect on security behavior. In particular, participants with a yearly income equivalent[3] to at least US $60K per year are estimated to score an average of at least 3.401 points higher than participants with income lower than (the equivalent of US) $60K per year.

### *Knowledge About Security*

The `Knowledge` explanatory variables also had a significant effect on participants' exhibited behavior. For example, scoring one point higher on the security test is estimated to increase the mean RSeBIS score by 0.265. Most interestingly

[3]Where "equivalent" is defined as being equal after correction by the country's Atlas method value, as discussed in our methodology section.

to us, the effect of a one-point increase in the knowledge test is about four times *lower* than a one-point increase in the scale measuring self-confidence security knowledge (despite self-confidence having a larger scale). In other words, self-confidence in computer-security knowledge has a larger effect on users' security behavior-intentions than their actual knowledge in computer security. Our results complement previous results showing than more knowledge about computer security does not necessary lead people to behave much more securely in reality, or that it is not the main factor driving actual computer-security behavior [17, 26, 54]. This raises the question of whether there are better alternatives to educating people about computer security in order to make them behave more securely.

Our initial analysis showed that technical people (people with degrees or jobs in technical areas) did not score significantly higher than non-technical people in the security test. This may be due to the fact that technical people are not necessarily security experts. By examining the regression's coefficients, however, one can learn that technical people are likely to behave more securely than non-technical people. A possible explanation for this is that technical people are instructed to follow secure behavior as part of their positions.

## DISCUSSION

In this section, we discuss the limitations of our design and methodology. Then, we turn to a discussion of our findings, and what we see as possible ways forward and take-aways for our research.

### *Limitations*

Our results should be interpreted in the context of several study limitations.

The use of nationality as a proxy for culture may be a coarse approximation. By doing so, we ignore the potential effects of expatriates, mixed national demographics, and shared experience due to technology. However, due to the drawbacks of possible alternatives (as explained in previous sections), using nationality to approximate culture remains one of the best available options [50, 52].

To reach a large number of participants from several countries, we use surveys to measure security behavior, knowledge, and confidence in this knowledge. Similarly to many studies that rely on surveys, our approach may suffer from response bias [19]. For instance, participants may prefer not to answer the survey questions honestly, and report more socially acceptable behavior, instead. In our survey, RSeBIS and the questions on self-confidence in security knowledge are the only questions potentially vulnerable to response bias. RSeBIS is closely tied to SeBIS, whose questions were shown not to correlate with social-desirability scales [16]. Moreover, we found that self-confidence has a significant effect on security behavior. This finding would have been unlikely had participants' responses been affected by response bias. Therefore, we believe that the potential effect of response bias on our results is either small or non-existent.

### *Survey Translation*

In this work, we developed a survey, RSeBIS, in English and followed a rigorous process to translate it into six other languages. Similarly to previous work [8, 51, 55], we found that reverse-worded questions were the main threat to the validity and reliability of the scales measured by our survey. While such questions may help avoid participants' response bias in some cases, they may lead cross-cultural studies to fail. Our study presents another evidence to support advice for avoiding reverse-worded questions in cross-cultural studies.

RSeBIS can be readily used by anyone interested in measuring security behavior-intentions among speakers of the seven languages that we considered in this paper. For example, researcher can use it to measure the intended security behavior of participants in their studies, while practitioners can use it to examine common behavior of employees in their organizations. (Egelman and Peer list other important use-cases [16].) Moreover, our translation methodology proved to produce reliable and valid scales in the languages considered. Thus, we believe it can be used to translate the scale to additional languages in the future.

### *Globalization of Usable Security Research*

Prior work in the area of usable privacy and security has focused mainly on western participants. Despite their importance, these samples represent a limited fraction of internet users. In this paper, we show such usable security studies may not generalize well to users from less studied regions. Thus, previous indicators and systems proposed to make people more secure (e.g., HTTPS indicators in browsers) may not be as effective on a global scale as previously thought. Our findings motivate future usable security research, by highlighting the need to conduct more global studies to design or tailor human-centered defenses.

### *Personalized Security Tools*

Previous research has shown that personalization of websites and ads can lead users to spend more time on websites and purchase more products [39, 48]. A similar approach may be viable to enhance the security behavior of users. Personalized security tools and policies may help foster more secure user behavior. Our work shows that users from certain cultures are likely to adopt riskier postures than users from other cultures. Therefore, we believe that future research should study the personalization of security with a specific focus on using users' culture, nationality, or location toward achieving that goal. For example, password creation policies or guidelines could be tailored to users' locations. As another example, to increase warning compliance, browsers could use different indicators to inform users in different locations that they are visiting malicious websites. In a similar vein, information about users' location could help set default system security settings [43].

### *Improving Security Behavior*

Our results also shed light on how knowledge of computer security affects actual behavior. Although people who know about computer security are likely to exhibit more secure behavior, the estimated effect size of knowledge on behavior is small. This finding indicates that to enhance users' secure behavior, improving user knowledge about security should *not*

be the sole focus, but rather should be one aspect among many. Other aspects, such as developing users' confidence in their ability to secure their devices and data—possibly by following a positive approach to teach them about secure behavior [44], may be at least as effective.

## CONCLUSION

We designed and ran an online study to explore how culture affects computer security behavior, interactively with other factors. Using an iterative design over 5,370 users (1,654 for our initial, negative validation, and 3,716 for subsequent hypothesis testing), we showed how to accomplish a robust translation of the survey we developed into other languages while preserving the reliability and validity of the scales it measures. We then collected 3,500 responses to our survey from participants located in seven different countries and analyzed these responses using linear regressions. Our results shed light on the differences in security behavior between participants from different cultures. For example, French participants exhibited the most secure behavior, whereas participants from Asian countries behaved less securely. Our results also show that self-confidence in computer security knowledge has a larger positive effect on security behavior compared to actual knowledge about computer security. Our findings motivate future research in usable security to be conducted more globally. They also motivate future work that studies customization of security tools based on people's culture or nationality. Additionally, we recommend looking into directions beyond user education to promote more secure behavior.

## ACKNOWLEDGMENTS

We would like to thank Aaron Harlap, Soo-Jin Moon, and Rui Xu for helping us translate RSeBIS into Russian, Korean, and Chinese. We are also grateful to Serge Egelman and Eyal Peer for sharing broad statistics on the demographics of participants from their study.

## REFERENCES

1. Galit Ailon. 2008. Mirror, mirror on the wall: Culture's consequences in a value test of its own design. *Academy of management review* 33, 4 (2008), 885–904.
2. Hashem Abdullah A Almakrami. 2015. *Online self-disclosure across cultures: A study of Facebook use in Saudi Arabia and Australia*. Ph.D. Dissertation. Queensland University of Technology.
3. Annie I Antón, Julia B Earp, and Jessica D Young. 2010. How internet users' privacy concerns have evolved since 2002. *IEEE Security & Privacy* 8, 1 (2010), 21–27.
4. Steven Bellman, Eric J Johnson, Stephen J Kobrin, and Gerald L Lohse. 2004. International differences in information privacy concerns: A global survey of consumers. *The Information Society* 20, 5 (2004), 313–324.
5. Sunil Chaudhary, Yan Zhao, Eleni Berki, Juri Valtanen, Linfeng Li, Marko Helenius, and Stylianos Mystakidis. 2015. A cross-cultural and gender-based perspective for online security: Exploring knowledge, skills and attitudes of higher education students. *IADIS International Journal on WWW/Internet* 13, 1 (2015).
6. China Internet Network Information Center. 2014. Statistical Survey on Internet Developmentin China2015.12. http://www.internetstatistik.se/wordpress/wp-content/uploads/2014/07/CNNIC_report.pdf. (2014). [Online; accessed 9/5/2016].
7. Hichang Cho, Milagros Rivera-Sánchez, and Sun Sun Lim. 2009. A multinational study on online privacy: global concerns and local responses. *New media & society* 11, 3 (2009), 395–416.
8. Robert Colosi. 2005. Negatively worded questions cause respondent confusion. *Proceedings of the Survey Research Methods Section, American Statistical Association (2005)* (2005), 2896–2903.
9. ComScore, Inc. 2015. 2015 Japan Digital Audience Report. https://www.comscore.com/layout/set/popup/Request/Presentations/2015/2015-Japan-Digital-Audience-Report?req=slides&pre=2015+Japan+Digital+Audience+Report. (2015). [Online; accessed 9/5/2016].
10. CRÉDOC. 2014. La diffusion des technologies de l'information et de la communication dans la société française (2014). http://www.arcep.fr/uploads/tx_gspublication/etude-CREDOC-diffusion-TIC-2014.pdf. (2014). [Online; accessed 9/5/2016].
11. Cross Marketing Inc. 2015. Online Research. http://global.cross-m.co.jp/solution/online/index.html. (2015). [Online; accessed 9/5/2016].
12. Dan Cvrcek, Marek Kumpost, Vashek Matyas, and George Danezis. 2006. A study on the value of location privacy. In *Proceedings of WPES*. DOI: http://dx.doi.org/10.1145/1179601.1179621
13. Jacqueline K Eastman, Bill Fredenberger, David Campbell, and Stephen Calvert. 1997. The Relationship between Status Consumption and Materialism: A Cross-Cultural Comparison of Chinese, Mexican, and American Student. *Journal of Marketing Theory and Practice* 5, 1 (1997), 52–66.
14. Serge Egelman, Marian Harbach, and Eyal Peer. 2016. Behavior Ever Follows Intention?: A Validation of the Security Behavior Intentions Scale (SeBIS). In *Proceedings of CHI*. DOI: http://dx.doi.org/10.1145/2858036.2858265
15. Serge Egelman and Eyal Peer. 2015a. The myth of the average user: Improving privacy and security systems through individualization. In *Proceedings of NSPW*. DOI: http://dx.doi.org/10.1145/2841113.2841115
16. Serge Egelman and Eyal Peer. 2015b. Scaling the Security Wall: Developing a Security Behavior Intentions Scale (SeBIS). In *Proceedings of ACM CHI*. DOI: http://dx.doi.org/10.1145/2702123.2702249

17. Alain Forget, Sarah Pearman, Jeremy Thomas, Alessandro Acquisti, Nicolas Christin, Lorrie Faith Cranor, Serge Egelman, Marian Harbach, and Rahul Telang. 2016. Do or Do Not, There Is No Try: User Engagement May Not Improve Security Outcomes. In *Proceedings of SOUPS*.
18. Tsutomu Fujinami and Christine Nanz. 2013. The 101 translation problems between Japanese and German/English. (2013).
19. Adrian Furnham. 1986. Response bias, social desirability and dissimulation. *Personality and individual differences* 7, 3 (1986), 385–400.
20. Gengo, Inc. 2016. Gengo. https://www.gengo.com. (2016). [Online; accessed 09/20/2016].
21. Güliz Ger and Russell W Belk. 1996. Cross-cultural differences in materialism. *Journal of economic psychology* 17, 1 (1996), 55–77.
22. Marian Harbach, Alexander De Luca, Nathan Malkin, and Serge Egelman. 2016. Keep on Lockin'in the Free World: A Multi-National Comparison of Smartphone Locking. In *Proceedings of ACM CHI*. DOI: http://dx.doi.org/10.1145/2858036.2858273
23. Geert Hofstede. 1984. *Culture's consequences: International differences in work-related values*. Vol. 5. Sage.
24. Geert Hofstede. 2001. *Culture's consequences: Comparing values, behaviors, institutions and organizations across nations*. Sage.
25. 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.
26. Iulia Ion, Rob Reeder, and Sunny Consolvo. 2015. "... no one can hack my mind": Comparing Expert and Non-Expert Security Practices. In *Proceedings of SOUPS*.
27. Iulia Ion, Niharika Sachdeva, Ponnurangam Kumaraguru, and Srdjan Čapkun. 2011. Home is safer than the cloud!: privacy concerns for consumer cloud storage. In *Proceedings of SOUPS*.
28. Ipsos. 2014. Ipsos Online Audience Measurement in The Arab World. http://fac.ksu.edu.sa/sites/default/files/online-audience-measurement.pdf. (2014). [Online; accessed 9/5/2016].
29. Ruogu Kang, Stephanie Brown, Laura Dabbish, and Sara B Kiesler. 2014. Privacy Attitudes of Mechanical Turk Workers and the US Public. In *Proceedings of SOUPS*.
30. Kristiina Karvonen, Lucas Cardholm, and Stefan Karlsson. 2000. Cultures of trust: A cross-cultural study on the formation of trust in an electronic environment. In *Proceedings of the nordic workshop on secure IT systems*.
31. Korean Internet & Security Agency. 2016. 2015 Survey on the Internet Usage Executive Summary. http://isis.kisa.or.kr/board/fileDown.jsp&pageId=060300&bbsId=10&itemId=345&athSeq=1. (2016). [Online; accessed 9/5/2016].
32. Hanna Krasnova and Natasha F Veltri. 2010. Privacy calculus on social networking sites: Explorative evidence from Germany and USA. In *Proceedings of HICSS*.
33. Lydia Kraus, Ina Wechsung, and Sebastian Möller. 2014. A comparison of privacy and security knowledge and privacy concern as influencing factors for mobile protection behavior. In *Proceedings of PPS*.
34. Ponnurangam Kumaraguru, Lorrie Faith Cranor, and Elaine Newton. 2005. Privacy perceptions in india and the united states: An interview study. In *Proceedings of TPRC*.
35. Jialiu Lin, Michael Benisch, Norman Sadeh, Jianwei Niu, Jason Hong, Banghui Lu, and Shaohui Guo. 2013. A comparative study of location-sharing privacy preferences in the United States and China. *Personal and ubiquitous computing* 17, 4 (2013), 697–711.
36. Macromill, Inc. 2015. Macromill. http://www.macromill.com/global/index.html. (2015). [Online; accessed 9/5/2016].
37. Brendan McSweeney. 2002. Hofstede's model of national cultural differences and their consequences: A triumph of faith-a failure of analysis. *Human relations* 55, 1 (2002), 89–118.
38. Ferenc Moksony. 1999. Small is beautiful. The use and interpretation of $R^2$ in social research. *Review of Sociology* (1999), 130–138.
39. Alan L Montgomery and Michael D Smith. 2009. Prospects for Personalization on the Internet. *Journal of Interactive Marketing* 23, 2 (2009), 130–137.
40. Richard G. Netemeyer, William O. Bearden, and Subhash Sharma. 2003. *Scaling procedures: Issues and applications*. Sage Publications.
41. Ofcom. 2015a. Gross national income per capita 2015, Atlas method and PPP. http://databank.worldbank.org/data/download/GNIPC.pdf. (2015). [Online; accessed 9/5/2016].
42. Ofcom. 2015b. International Communications Market Report 2015. http://stakeholders.ofcom.org.uk/binaries/research/cmr/cmr15/icmr15/icmr_2015.pdf. (2015). [Online; accessed 9/5/2016].
43. Mehrbod Sharifi, Eugene Fink, and Jaime G Carbonell. 2010. Learning of personalized security settings. (2010).
44. Mikko Siponen, M Adam Mahmood, and Seppo Pahnila. 2009. Technical opinion Are employees putting your company at risk by not following information security policies? *Commun. ACM* 52, 12 (2009), 145–147. DOI: http://dx.doi.org/10.1145/1610252.1610289

45. Tom W Smith. 2003. Developing comparable questions in cross-national surveys. *Cross-cultural survey methods* (2003), 69–92.

46. Statistics Bureau. 2011. 2010 National Census in Japan (in Japanese). http://www.stat.go.jp/data/kokusei/2010/index.htm. (2011). [Online; accessed 5/10/2016].

47. Mohsen Tavakol and Reg Dennick. 2011. Making sense of Cronbach's alpha. *International journal of medical education* 2 (2011), 53.

48. Narongsak Thongpapanl and Abdul Rehman Ashraf. 2011. Enhancing online performance through website content and personalization. *Journal of computer information systems* 52, 1 (2011), 3–13.

49. TNS Russia. 2015. Web Index. http://tns-global.ru/services/media/media-audience/internet/information/. (2015). [Online; accessed 9/5/2016].

50. Blase Ur and Yang Wang. 2013. A cross-cultural framework for protecting user privacy in online social media. In *Proceedings of WWW*. DOI: http://dx.doi.org/10.1145/2487788.2488037

51. Eric Van Sonderen, Robbert Sanderman, and James C Coyne. 2013. Ineffectiveness of reverse wording of questionnaire items: Let's learn from cows in the rain. *PloS one* 8, 7 (2013), e68967.

52. Yang Wang, Gregory Norcie, and Lorrie Faith Cranor. 2011. Who is concerned about what? A study of American, Chinese and Indian users' privacy concerns on social network sites. In *Proceedings of TRUST*.

53. Yang Wang, Huichuan Xia, and Yun Huang. 2016. Examining American and Chinese Internet Users' Contextual Privacy Preferences of Behavioral Advertising. In *Proceedings of CSCW*. DOI: http://dx.doi.org/10.1145/2818048.2819941

54. Rick Wash and Emilee Rader. 2015. Too much knowledge? security beliefs and protective behaviors among united states internet users. In *Proceedings of SOUPS*.

55. Nancy Wong, Aric Rindfleisch, and James E Burroughs. 2003. Do reverse-worded items confound measures in cross-cultural consumer research? The case of the material values scale. *Journal of consumer research* 30, 1 (2003), 72–91.

56. World Bank. 2016. World Bank Atlas Method. (2016). http://go.worldbank.org/IEH2RL06U0 [Online; accessed 12/31/2016].

57. Chen Zhao, Pamela Hinds, and Ge Gao. 2012. How and to whom people share: the role of culture in self-disclosure in online communities. In *Proceedings of CSCW*. DOI:http://dx.doi.org/10.1145/2145204.2145219

## APPENDIX

**RSeBIS** (5-point Likert scale; from "never" to "always"):

1. I set my computer screen to automatically lock if I don't use it for a prolonged period of time.
2. I use a password/passcode to unlock my laptop or tablet.
3. I manually lock my computer screen when I step away from it.
4. I use a PIN or passcode to unlock my mobile phone.
5. I change my passwords even if it is not needed.
6. I use different passwords for different accounts that I have.
7. When I create a new online account, I try to use a password that goes beyond the site's minimum requirements.
8. I include special characters in my password even if it's not required.
9. When someone sends me a link, I open it only after verifying where it goes.
10. I know what website I'm visiting by looking at the URL bar, rather than by the website's look and feel.
11. I verify that information will be sent securely (e.g., SSL, "https://", a lock icon) before I submit it to websites.
12. When browsing websites, I mouseover links to see where they go, before clicking them.
13. If I discover a security problem, I fix or report it rather than assuming somebody else will.
14. When I'm prompted about a software update, I install it right away.
15. I try to make sure that the programs I use are up-to-date.
16. I verify that my anti-virus software has been regularly updating itself.

**Test of Security Knowledge** (true/false):

1. My Internet provider and location can be disclosed from my IP address.
2. My telephone number can be disclosed from my IP addresses.
3. The web browser information of my device can be disclosed to the operators of websites.
4. Since Wi-Fi networks in coffee shops are secured by the coffee shop owners, I can use them to send sensitive data such as credit card information.
5. Password comprised of random characters are harder for attackers to guess than passwords comprised of common words and phrases.
6. If I receive an email that tells me to change my password, and links me to the web page, I should change my password immediately.
7. My devices are safe from being infected while browsing the web because web browsers only display information.
8. It is impossible to confirm whether secure communication is being used between my device and a website.
9. My information can be stolen if a website that I visit masquerades as a famous website (e.g., amazon.com).

10. I may suffer from monetary loss if a website that I visit masquerades as a famous website.
11. My devices and accounts may be put at risk if I make a typing mistake while entering the address of a website.
12. My IP address is secret and it is unsafe to share it with anyone.
13. If my web browser does not show a green lock when I visit a website, then I can deduce that the website it is malicious.
14. It is safe to open links that appear in emails in my inbox.
15. It is safe to open attachments received via email.
16. I use private browsing mode to protect my machine from being infected.
17. It is safe to use anti-virus software downloaded through P2P file sharing services.
18. Machines are safe from infections unless users actively download malware.

**Self-confidence in Security Knowledge** (5-point Likert scale; from "strongly disagree to "strongly agree"):

1. I know about countermeasures for keeping the data on my device from being exploited.
2. I know about countermeasures to protect myself from monetary loss when using the Internet.
3. I know about countermeasures to prevent my IDs or Passwords being stolen.
4. I know about countermeasures to prevent my devices from being compromised.
5. I know about countermeasures to protect me from being deceived by fake web sites.
6. I know about countermeasures to prevent my data from being stolen during web browsing.