# The Study of Cybersecurity Self-Efficacy: A Systematic Literature Review of Methodology

Borgert, Nele
*Ruhr University Bochum*

Friedauer, Jennifer
*Ruhr University Bochum*

Böse, Imke
*Ruhr University Bochum*

Sasse, M. Angela
*Ruhr University Bochum*

Elson, Malte
*Ruhr University Bochum*

## Abstract

The aim of this project is to obtain a systematic assessment of the methodological practices in Cybersecurity Self-Efficacy (CSE) research. CSE is the belief about one's own ability to enact skills related to IT security or privacy, a psychological disposition that directly affects security behaviors. The co-determining effect of self-efficacy beliefs on motivation, and consequently performing behaviors, works through impacting mechanisms on goals, outcome expectations, and sociostructural factors. Implementing effective security behaviors is a major key to maintaining safe use of today's smart technologies and ensuring one's own data is protected. Given the sensitivity of data collected by e.g., smart home devices, its protection is of high relevance. Still, there remains a lack of systematic research that deals with the arisen ambiguity of CSE. We conducted a systematic literature review of general methodology and psychometrics, which identified a total of 1,769 potentially relevant research papers on the CSE topic.

## 1 Introduction

### 1.1 Background

Over the past few years, there has been a rapid integration of smart devices (e.g., Amazon Alexa) throughout our day-to-day lives. The popularity of these products among consumers can easily be attributed to the convenience they are designed to provide [9]. However, concerns have been raised about the security and privacy of technologies, as they are strongly shaped by the users' behavior [10], meaning that the protection of individual data depends on each individual consumer in ways that are often unknown to inexperienced users. This risk of vulnerable IT security and data protection is not transparent for many consumers and yet it is of utmost importance due to the strong sensitivity of the processed data. A high sensitivity of data is especially relevant in the context of smart homes, since appliances often have access to video or audio recordings in core privacy spaces [5].

Consequently, implementing effective security behaviors is key to maintaining safe use of smart devices and ensuring one's own data is protected. A motivational construct that inherently affects individual security behaviors is Cybersecurity Self-Efficacy (CSE). CSE is the belief about one's own ability to enact skills related to IT security or privacy (cf. [1]). The co-determining effect of self-efficacy beliefs on motivation, and consequently performing behaviors, works through impacting mechanisms on goals, outcome expectations, and sociostructural factors [2]. Due to the importance of CSE for security behaviors [3], it is a widespread psychological construct in literature on cybersecurity.

Though there were some prior attempts to synthesize the multiverse of research strings on self-efficacy regarding behavioral IT security by formulating a holistic definition and providing measurement dimensions [7], there still remains a lack of systematic approaches that address the resulted ambiguity of key criteria of CSE and identify current science trends in the field. To this end, we have conducted a systematic literature review on CSE methodology.

### 1.2 Research Questions

Our review aims to systematically investigate published empirical research on CSE. We focus the research questions on three essential areas of interest: general methodology, theoretical models, and manipulating interventions. Reviewing general methodology is motivated by apparent divergent measurements, which jeopardizes study comparability across disciplines and the validity of inferences based on study results. It is difficult to buildup cumulative evidence when existing


*USENIX Symposium on Usable Privacy and Security (SOUPS)* 2021.
August 8–10, 2021, Vancouver, B.C., Canada.

methodological strengths and shortcomings are obscure. For theoretical models, it is unclear what role CSE plays due to the field's definitional ambiguity of self-efficacy. The context in which CSE variables are studied might differ and hence, the antecedents and consequences of CSE are in need of aggregation to paint a comprehensive picture as well as allow actionable implications for practitioners. This is also a prerequisite to grounded interventions that are designed to manipulate CSE. Their evidence based relevance for practice advises is undisputed. Finally, all those essential review areas of interest apply likewise but distinctly to human interactions with smart home systems. The population of smart home users, their CSE and its role, as well as paradigms for interventions might be different partly due to the technology's appeal (new prevalence and primary tasks of convenience), which motivated the sub-grouped review approach. Thus, our literature research questions are as follows:

1. What are the demographics of studies of cybersecurity self-efficacy?
2. What measures are used to assess cybersecurity self-efficacy? What are the scale characteristics and reported psychometric properties?
3. What role does cybersecurity self-efficacy play in the theoretical or research models of the studies?
4. Do the studies report interventions that have been carried out to manipulate cybersecurity self-efficacy? If so, what was their approach?
5. Are there studies of cybersecurity self-efficacy regarding smart home devices? If so: what is their respective answer to questions 1-4

## 2 Methods

Our systematic literature review has been preregistered at Open Science Framework (https://osf.io/vf6bn/?view_only=8c90aa7745634002b44310d460b731cd) before data was collected. In addition, we followed the international PROSPERO scheme for documentation standards of systematic review protocols for research with human subjects.

### 2.1 Search Strategy

The search strategy incorporated 18 electronic databases to fairly embrace the interdisciplinary nature of CSE. We queried EBSCOhost (which includes the databases Academic Search Premier, APA PsycArticles, APA PsycInfo, Historical Abstracts, OpenDissertations, PSYNDEX Literature with PSYNDEX Tests), IEEE Xplore, ACM Digital Library, Science Direct, dimensions.ai, arXiv, Scopus, Web of Science (i.e. WOS, KJD, MEDLINE, RSCI, SCIELO), and the Wiley Online Library.

To promote a comprehensive and reproducible search that reduces potential biases by our research team, we used group discussions as well as a quasi-automated text mining and keyword co-occurrence network method by [6] (R version: 4.0.3; litsearchr package version: 1.0.0). Our final search terms had the following syntax:

```
"self-efficacy" AND ("cybersecurity" OR
"cyber security" OR "information security"
OR "IT security" OR "information technology
security" OR "IS security" OR "information
system security" OR "wireless security" OR
"home wireless security" OR "usable security"
OR "computer security" OR "data protection" OR
"data security" OR "personal data" OR "privacy"
OR "security threat" OR "wireless network" OR
"device security").
```

### 2.2 Study Selection

The applied selection criteria covered the following aspects: studies should be empirical and have a sample size $N > 1$, be published after 2009 and written in English, measure self-efficacy regarding IT security or privacy, and examine the relationship between self-efficacy and IT security or privacy. All study types, populations, cultures, countries, interventions, or exposures were included.

Data was collected on March 18, 2021. The literature management software Citavi (Build Number: 6.4.0.35), automatically removed duplicates before documents were manually sifted. To ensure the relevance of our search string, records had to include the before-mentioned keywords in their abstracts. The EBSCOhost engine was not able to limit search results to both, the search in abstract texts and adherence to our search string with the exact phrases indicated by quotation marks. Thus, EBSCOhost produced a number of false positives that were then correctly reduced by an implemented python script. Next, identified studies were reviewed by three (blinded) persons regarding the selection criteria in three separate sifts: first based on title, then abstract, and finally full text information.

### 2.3 Coding Process

We defined all coded study characteristics in codebooks before the search was conducted. Prior to the process of study selection and data extraction, reviewers were trained until interrater agreement reached a satisfactory level (iota coefficient $> 0.6$ for the training data set) [8]. One key variable for each research question was pooled into iota indices for nominal and quantitative data [4]; $\iota = .647$ and $\iota = .992$ respectively (R version: 4.0.3; irr package version: 0.84.1).

To further combat potential biases and errors, we randomized (using randomizer.org) the study sample in two steps. First, before the title sift, studies were randomized and split

into three blocks, of which two random blocks were assigned to each reviewer. Second, we re-randomized the remaining sample after our full text sift the same way. Thus, each record was sifted and coded by two independent reviewers. Disagreements were discussed in group, where if necessary a 2/3 vote won.

## 3 Preliminary Results

Our search yielded a total of 1,769 records. After the automatic duplicates' removal, we had 1,017 titles in Citavi, which were again reduced to 862 titles due to the supplementary EBSCOhost python script. Overall, this resulted in 696 actual records to be screened in the sifts. As this is still an ongoing research project, the eventual sample size is not yet finalized. By the time of the conference, the poster will contain completed statistical results and a supporting discussion.

Eligible studies that were included in the synthesis seem to be published mainly in the second half of the last decade. A majority of the studies are surveys conducted in certain pre-established organizations (such as businesses, universities, schools, or camps) with sample sizes larger than 200 respondents. Sample recruitment mainly occurs in the USA, but still includes a variety of diverse countries.

Many studies refer to the same scales when introducing their own measures used to assess CSE. At first, this seemed to indicate a largely standardized approach. However, instruments are frequently modified by, for example, changing their items' wording, reducing the item number, or merging different scales. This manifests itself also in the almost non-existent validating test development. Yet, statistics for the internal consistency are on average satisfactory.

CSE's role in current research models is mostly portrait as a causal acting characteristic of persons. In consequence, studies predominantly survey IT related consequences of CSE, such as specific attitudes, expectations, intentions, and security behaviors. This entails the fact that merely a few studies include designs with interventions that are implemented to affect CSE. Only about a tenth of studies try to enhance the motivational aspect of self-efficacy in this domain. Nevertheless, appropriate CSE interventions would be of high interest as we could observe, inter alia, behavioral changes in relevant performance variables for IT security or privacy.

IT security and privacy of smart home devices in particular plays only a marginal role. So far, our review identified one study on smart homes, wherefore at the present stage, there is no opportunity to synthesize the characteristics of interest.

## 4 Conclusion

In the present literature review, we investigate the psychological construct of self-efficacy in IT security and privacy. CSE is visibly important to the human-computer interaction community as the demographics of studies demonstrate. We try to synthesize identified literature records in three fundamental areas: general methodology, theoretical models, and manipulating interventions. All of these areas are essential for both future research efforts and practical indications.

The popularity and dispersion of CSE research might be due to its assumed inherent influence on security intentions and behaviors, which themselves play an increasingly important role in today's technologized world. Research models captured by this review confirm this assumption. Still, our results paint a deeply fragmented picture of model approaches.

The lack of CSE measurements developed in a psychometrical fashion for its own purpose of validation research, and hence, the staggering number of ad-hoc self-developed scales is worrisome with regard to study robustness and replicability. Here, we recognize the need of high quality instruments that are efficient and usable for a multitude of research strings.

## Acknowledgments

This work was supported by the German Federal Ministry of Education and Research (BMBF) (grant no. V5DISO056-03). M.E. is supported by the Digital Society research program funded by the Ministry of Culture and Science of North Rhine-Westphalia, Germany (grant no. 1706dgn006).

## References

[1] Albert Bandura. Self-efficacy. In Irving B. Weiner and W. Edward Craighead, editors, *The Corsini Encyclopedia of Psychology*. John Wiley & Sons, Inc, Hoboken, NJ, USA, 2010.

[2] Albert Bandura. On the functional properties of perceived self-efficacy revisited. *Journal of Management*, 38(1):9–44, 2012.

[3] Mark Chan, Irene Woon, and Atreyi Kankanhalli. Perceptions of information security in the workplace: Linking information security climate to compliant behavior. *Journal of Information Privacy and Security*, 1(3):18–41, 2005.

[4] Matthias Gamer, Jim Lemon, Ian Fellows, and Puspendra Singh. Various coefficients of interrater reliability and agreemen: Package 'irr', 2019.

[5] Dan Goodin. Home alarm tech backdoored security cameras to spy on customers having sex: Employee for adt accessed ˜200 customer cams on more than 9,600 occasions. *ars technica*, 22.01.2021.

[6] Eliza M. Grames, Andrew N. Stillman, Morgan W. Tingley, and Chris S. Elphick. An automated approach to

identifying search terms for systematic reviews using keyword co–occurrence networks. *Methods in Ecology and Evolution*, 10(10):1645–1654, 2019.

[7] Wu He, Xiaohong Yuan, and Xin Tian. The self-efficacy variable in behavioral information security research. In *2014 Enterprise Systems Conference*, pages 28–32.

[8] Harald Janson and Ulf Olsson. A measure of agreement for interval or nominal multivariate observations. *Educational and Psychological Measurement*, 61(2):277–289, 2001.

[9] Eunil Park, Yongwoo Cho, Jinyoung Han, and Sang Jib Kwon. Comprehensive approaches to user acceptance of internet of things in a smart home environment. *IEEE Internet of Things Journal*, 4(6):2342–2350, 2017.

[10] Hyeun-Suk Rhee, Cheongtag Kim, and Young U. Ryu. Self-efficacy in information security: Its influence on end users' information security practice behavior. *Computers & Security*, 28(8):816–826, 2009.