A systematic review of scales for measuring information security culture

Špela Orehek and Gregor Petrič
Faculty of Social Sciences, Center for Methodology and Informatics, University of Ljubljana, Ljubljana, Slovenia

Received 13 December 2019
Revised 23 March 2020
25 May 2020
29 July 2020
1 September 2020
Accepted 1 September 2020

Abstract

Purpose – The concept of information security culture, which recently gained increased attention, aims to comprehensively grasp socio-cultural mechanisms that have an impact on organizational security. Different measurement instruments have been developed to measure and assess information security culture using survey-based tools. However, the content, breadth and face validity of these scales vary greatly. This study aims to identify and provide an overview of the scales that are used to measure information security culture and to evaluate the rigor of reported scale development and validation procedures.

Design/methodology/approach – Papers that introduce a new or adapt an existing scale of information security culture were systematically reviewed to evaluate scales of information security culture. A standard search strategy was applied to identify 19 relevant scales, which were evaluated based on the framework of 16 criteria pertaining to the rigor of reported operationalization and the reported validity and reliability of the identified scales.

Findings – The results show that the rigor with which scales of information security culture are validated varies greatly and that none of the scales meet all the evaluation criteria. Moreover, most of the studies provide somewhat limited evidence of the validation of scales, indicating room for further improvement. Particularly, critical issues seem to be the lack of evidence regarding discriminant and criterion validity and incomplete documentation of the operationalization process.

Research limitations/implications – Researchers focusing on the human factor in information security need to reach a certain level of agreement on the essential elements of the concept of information security culture. Future studies need to build on existing scales, address their limitations and gain further evidence regarding the validity of scales of information security culture. Further research should also investigate the quality of definitions and make expert assessments of the content fit between concepts and items.

Practical implications – Organizations that aim to assess the level of information security culture among employees can use the results of this systematic review to support the selection of an adequate measurement scale. However, caution is needed for scales that provide limited evidence of validation.

Originality/value – This is the first study that offers a critical evaluation of existing scales of information security culture. The results have decision-making value for researchers who intend to conduct survey-based examinations of information security culture.

Keywords Information security culture, Information security, Measurement, Scales, Validity, Systematic review, Surveys, Assessments, Methodology, Measurement, Meta-analysis

Paper type Research paper

This work was supported by the Slovenian Research Agency within the "Young researchers" program [grant number P5-0168].

Information & Computer Security
Vol. 29 No. 1, 2021
pp. 133-158
Emerald Publishing Limited
2056-4961
DOI 10.1108/ICS-12-2019-0140

1. Introduction

In contemporary societies, organizational processes are inextricably associated with the use of information and communications technologies (ICTs), which expose organizations to information security risks and threats. This is evidenced by the numerous security incidents in the past decade. Information security researchers and professionals are increasingly insisting that security risks and threats cannot be effectively avoided by solely technical means and that organizational human capital, which can influence the security of an entire organization, must be considered (Gordon and Loeb, 2005; Rančigaj and Lobnikar, 2012; Tsohou *et al.*, 2015). Organizations can implement the latest technological security solutions, but employees are still the ones who (often unknowingly) invite security breaches through careless behavior, which results from a poor information security culture (Singh *et al.*, 2014; Tsohou *et al.*, 2015). From the perspective that views ICTs as socio-technical systems (Kling, 2007), organizations can be secure only when both their technical and socio-cultural elements are in harmony.

1.1 The human factor in information security

The social scientific aspects of information security are commonly considered "the human factor" of information security and are researched under the umbrella of the behavioral information security approach. Different theoretical (mostly psychological) models have been applied in this field, and a number of concepts have emerged, such as information security awareness, conscious care behavior, compliance with security policies, information protection culture and cybersecurity culture. In particular, in recent years, the concept of information security culture has gained significant attention in both academia and industry. It aims to transcend predominant psychological perspectives and comprehensively address the human factor of information security in the organizational context by considering the cognitive, behavioral, attitudinal, normative and other aspects of employees that affect the establishment of information security (Al Hogail and Mirza, 2015; Da Veiga and Martins, 2015a; Nævestad *et al.*, 2018; Roer, 2015). Recent scientific articles and security reports (Budge *et al.*, 2018; European Union Agency for Network and Information Security [ENISA], 2017; Moody *et al.*, 2018) show that the concept of information security culture is becoming particularly important.

Information security culture generally refers to the formation of adequate beliefs and values regarding security that guide employees to establish a safe organizational information environment (Al Hogail and Mirza, 2015, p. 286). More precisely, Da Veiga and Eloff (2010, p. 198) define information security culture as "the attitudes, assumptions, beliefs, values and knowledge that employees/stakeholders use to interact with the organization's systems and procedures at any point in time." In other words, employees have certain values, and therefore, demonstrate certain behaviors that can either support the protection of organizational assets or endanger them (Al Hogail and Mirza, 2015). Organizations can have certain regulations and guidelines about employee conduct that relate to safeguarding information systems, but employees can choose whether to behave in line with those regulations (Martins and Eloff, 2002). Thus, if the information security culture is weak, it may inhibit the security of organizations (Da Veiga, 2016). When a positive information security culture is developed, employees understand that ensuring security is an integral part of their job. Moreover, employees in such conditions actively practice good security habits and make security-minded decisions (Paulsen and Coulson, 2011).

Information security culture has a plethora of definitions and understandings in the scientific literature. It is understood as a subset of the wider concept of cybersecurity culture, which refers to human-induced security issues on a wider population level, while the

concept of information security culture is considered to be relevant within organizational contexts (Da Veiga, 2016). Moreover, we identify two different types of studies about how information security culture is conceptualized and related to other phenomena. On the one hand, there are studies that consider information security culture as one of the independent variables that is relevant to explaining organizational human behavior. In such studies, it is usually measured as a unidimensional concept with several items. On the other hand, there are studies that see information security culture as an overarching construct that includes other relevant concepts in the field of security behavior research. Within the latter studies, information security culture is viewed as a multidimensional concept that contains awareness, knowledge, attitudes, behaviors, intentions, beliefs, values and other relevant concepts, resulting in multidimensional measurement instruments with large numbers of items.

Along with becoming an important theoretical concept in the field of information security, information security culture carries important practical consequences. Organizations must develop a certain level of information security culture to reduce or minimize the security risks incurred by employees when they use organizational ICTs (Da Veiga and Eloff, 2010) or personal technologies in an organizational context. One of the main goals of promoting information security culture is to create and sustain an understanding of information security as an issue pertaining to each employee and their responsibilities and conduct in an organization (Al Hogail, 2015; Schlienger and Teufel, 2003). Knowledge of what organizational security culture is and how to manage it has become vital for the management of contemporary information-technology-dependent organizations (Furnell and Thomson, 2009; Tsohou *et al.*, 2015).

1.2 Necessity of valid and reliable measurements of information security culture

With the emergence of the concept of information security culture, a need to measure this phenomenon quickly developed and a number of measurement scales [1] emerged in the past decade. Measuring information security culture is important not only for scientific research of information security but also for practical reasons. When an organization measures its information security culture, it receives feedback on its employees' security behavior, their perception of the importance of security and peer behavior, their knowledge, values and other aspects of information security culture. In this way, measuring information security culture can help an organization to identify its weak links and develop guidelines to resolve them (or identify strong links and reaffirm them). The European Union Agency for Network and Information Security specifically calls for measurement of information security culture in organizations to assess and improve the human factor of security (European Union Agency for Network and Information Security [ENISA], 2017). Measuring complex theoretical concepts such as information security culture in a valid and reliable way is difficult, but not impossible.

However, at face value, existing measurement instruments greatly vary in terms of the rigor of the operationalization process, content of the items and evidence regarding the quality of measures. Thus, a critical overview of scales for measuring information security culture is urgently needed to provide grounds for the future development of high-quality measurement instruments and enable valid insights and reliable predictions. The importance of sound validation of scales is understood as one of the most important elements in research (Schoenfeldt, 1984), especially in the context of information security, as weakly validated scales might lead to a wrong assessment of information security culture in an organization and, in turn, to decisions with devastating consequences.

Therefore, this study aims to identify and provide an overview of the scales that are used to measure information security culture and to evaluate the rigor of reported scale

development and validation procedures. More precisely, we seek to answer three research questions:

RQ1. Which measurement scales are used to measure information security culture?

RQ2. What is the rigor of the reported operationalization of the identified scales?

RQ3. What is the reported validity and reliability of the identified scales?

2. Method

Systematic review is a commonly used method to obtain an overview of a certain scale and to collect and analyze data from a series of studies. It enables critical evaluation of identified scales, and the findings contribute to further optimization of those scales (Moher *et al.*, 2009).

2.1 Search strategy

To answer the first research question, we used a search strategy according to the preferred reporting items for systematic reviews and meta-analyzes procedure (Liberati *et al.*, 2009). Then, we assessed the results according to predefined methodological criteria to address the second and third research questions. The primary selection criteria for the inclusion of studies were as follows:

- The study is published as a scientific paper.
- The study involves empirical research with survey-based measure (s) of information security culture.

Identification: The search was conducted within the Scopus and Web of Science databases, which contain peer-reviewed scientific journals, books and conference proceedings. The search was conducted on June 27, 2019. The following Boolean search was performed: "TITLE-ABS-KEY ("information security culture" OR "culture of information security" OR "cybersecurity culture" OR "security culture") AND (survey OR measure OR scale OR questionnaire OR model)" [2]. The search procedure (based on title, abstract and keywords) yielded 441 hits. In total, 92 duplicates were found and excluded from further analysis.

Screening: A total of 349 hits were screened for general adequacy in terms of whether they addressed information security culture. Screening was based on the examination of titles and abstracts. During this phase, 114 sources were eliminated, which meant that 235 relevant sources remained according to the selection criteria listed above.

Eligibility: The last phase of the search strategy involved a thorough analysis of the studies that were selected in the screening phase. All relevant identified articles were accessed using a digital library service provided by the University of Ljubljana or were obtained via open access [3]. In particular, we focused on information about survey-based measurement of information security culture. Out of the 235 articles, 216 were excluded for various reasons: irrelevance of the topic (58 papers), theoretical papers without empirical research (69 papers; those were not identified as non-empirical in the previous phase because of their uninformative abstracts) and use of qualitative methodology (52 papers). In addition, 14 papers were excluded because they used identical scales from one author. If the authors used the same measurement instrument for the same number of items and sample data, we combined all criteria together into one evaluation [4]. However, if the scales were not identical, the one with more reporting characteristics was included in our analysis. Further, 17 papers were excluded because their content did not refer to the methodological aspects of measuring instruments (i.e. the instruments

were not introduced or the reported characteristics were limited to descriptive statistics). Six papers were excluded because they were not written in English.

Inclusion: Ultimately, 19 studies were included in the systematic review (Figure 1).

2.2 Data extraction and criteria for evaluation

There are a number of works in the field of social science methodology that define the criteria for developing, validating and reporting measurement scales. In this paper, we propose an evaluation framework based on some of the most-cited sources related to the measurement of social science phenomena, which are often understood as providing standards for measurement. This body of literature includes, but is not limited to, guidelines established by the American Educational Research Association in cooperation with the

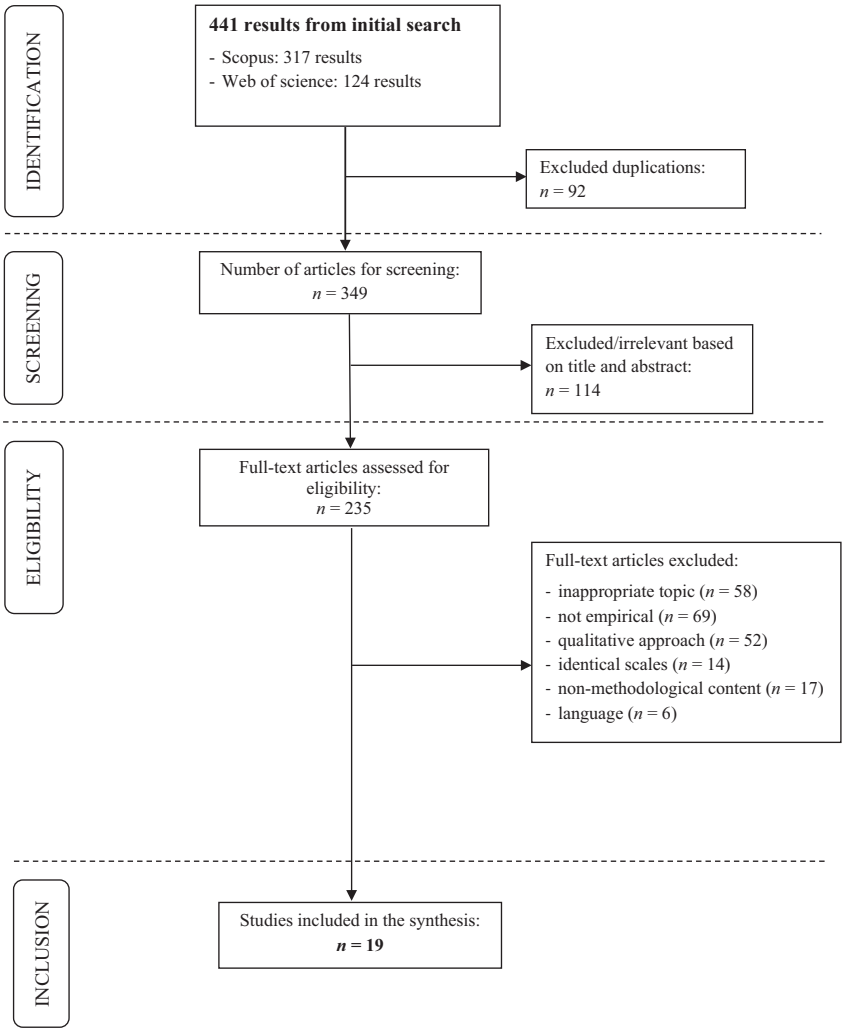

Figure 1.
Flow diagram of the search procedure based on the PRISMA procedure

American Psychological Association and National Council on Measurement in Education (AERA, 1992); the publication manual of the American Psychological Association (APA), 2020; and the work of DeVellis (2016). Our evaluation framework encompasses two main categories of methodological criteria, namely, the rigor of the reported operationalization of information security culture scales (Table 1) and the reported quality of measurement in terms of the reported empirical validity and reliability (Table 2). The analysis of the first set of criteria will serve to answer the second research question and the analysis of the second set of criteria will help us answer the third research question. The next paragraphs offer a detailed presentation of both categories of criteria.

The rigor of operationalization refers to the quality of reporting of the operationalization process in terms of how strictly authors follow the proposed guidelines for (documenting) operationalization and satisfy certain conditions in this process [5]. These steps are necessary but not sufficient to obtain a reliable and valid measurement instrument (Carmines and Zeller, 1979). The process of operationalization starts with an essential definition of a concept to be measured, identification of the concept's (possible) dimensions and their definitions, development and testing of an initial pool of items and establishment of a finalized scale with a clear set of items, introductory text and answer categories and basic statistics for the items (AERA, 1992; DeVellis, 2016; Neuman, 2014). In other words, the elements that are listed in Table 1 can be understood as necessary elements for a published study that introduces a new/adapted scale (Jackson, 2009; DeVellis, 2016; American Psychological Association [APA], 2020). In addition to coding whether quantitative elements of operationalization were absent or present in the published papers,

Table 1.
Criteria of the rigor of operationalization

Rigor of operationalization	Coded criteria	Empirical criteria/coding values
Conceptual background of items	Definition of concept	Reported (1); not reported (0)
	Definition of dimensions	Reported (1); not reported (0)
	Source of items	Reported (1); not reported (0)
Testing items	Expert review	Reported (1); not reported (0)
	Pilot test	Reported (1); not reported (0)
Sample information	Sample size	Higher than 300 or 5 times the number of items (1); lower or not reported (0)
	Response rate	Higher than 30% (1); lower or not reported (0)[a]
	Sample characteristics	Presence of basic sample characteristics (gender, age and education) (1); not reported (0) 1
Univariate and bivariate analysis	Descriptive statistics of the items or factors	Presence of basic descriptive statistics (mean, standard deviation, etc). for individual items or dimensions (1); not reported (0)
	Correlation between dimensions, factors or items	Presence of a correlation/covariance matrix or reported coefficients (1); not reported (0)
Measurement model fit assessment	Exploratory factor analysis	KMO test: >0.5; Bartlett's test: <0.05 (1), not reported or not fulfilled (0)[a]
	Confirmatory factor analysis[1]	χ^2/df < 3, RMSEA/RMR/SRMR: <0.1 and CFI/TLI/RNI/Bollen's delta 2: >0.9 (1); not reported (0)[a]

Note: [a]In case of partial fulfilled criteria, half point was assigned

we coded whether they satisfy the standards set by the methodological literature. All criteria pertaining to the rigor of operationalization and their coding values are listed in Table 1.

Regarding the conceptual background of operationalization, essential definitions are mandatory because they determine the content domain of a certain concept and serve as the starting point for the development of items (Bollen, 1989; DeVellis, 2016). Content validity can be assessed based on how well the definition of the concept is "translated" into the items (Trochim, 2006). According to the methodological literature, this is a step-by-step procedure that involves description of the source of items (i.e. the literature, concepts and qualitative interviews), (external) expert assessment of the fit between the definition of a concept and the content of items, and pilot testing of items (AERA, 1992; American Psychological Association [APA], 2020; DeVellis, 2016). In this evaluation, we do not evaluate the quality of definitions or the quality of pilot tests. Instead, we only report whether these elements are present and published in the paper (e.g. whether a definition of the concept is present or not, whether the scale was pilot tested or not).

The operationalization also includes requirements regarding the sample with which the scale is measured. We extracted information about the sample size and response rate, as these are important to obtain unbiased empirical parameters. Low response rates and an uncharacteristic sample structure can result in a biased sample and yield biased measurement results (Bollen, 1989; Hinkin, 1998). To statistically validate the measurement instrument, the sample size must be at least 300 units or five times larger than the number of items in a scale (Hair *et al.*, 2014; Nunnally and Bernstein, 1994; Tinsley and Tinsley, 1987). In addition, the response rate must be at least 30% in the case of online surveys to avoid a non-response bias (Callegaro *et al.*, 2015; Dillman *et al.*, 2014). In the event that a study collected data on multiple organizations and a sufficient response rate was achieved in only some cases, the criterion was coded as partially fulfilled. In addition, methodological guidelines suggest that studies should present the basic characteristics of the sample from which data for scale validation were collected. Sample characteristics were considered to be present if the main socio-demographic characteristics (i.e. gender, age and education) were reported and if one or two of these characteristics were reported, the criterion was considered to be partially fulfilled. The methodological literature also suggests that basic statistical information about items, such as descriptive statistics of items and/or dimensions and the correlation matrix between items or dimensions, should be provided (DeVellis, 2016; Jackson, 2009).

Finally, an important empirical step of operationalization is exploration and confirmation of the assumed measurement model (Thompson, 2004). Exploratory and confirmatory factor analyzes are required for establishing a valid new or adapted scale (Hinkin, 1998; Yong and Pearce, 2013). Factor analyzes themselves are not evidence of construct validity, yet they can

Element of quality of measuring	Coded criteria	Empirical criteria/coding values
Construct validity	Convergent validity	Factorial weights: >0.4 or AVE: >0.5
	Discriminant validity	Low correlations between factors or $\sqrt{\text{AVE}}$ greater than inter-construct correlations[a]
	Criterion validity	Correlation with criterion variable: >0.5
Reliability	Internal consistency	Cronbach's alpha: >0.7 or CR: >0.6[b]

Notes: [a]Other relevant tests for discriminant validity were also taken into account for evaluation, if they were reported. [b]In case of partial fulfilled criteria, half point was assigned

Table 2. Criteria of reported empirical validity and reliability

be understood as a technical precondition for establishing construct validity. We must differentiate between exploratory factor analysis (EFA), which is used for the initial exploration of a measurement instrument and confirmatory factor analysis (CFA), which reports whether the proposed measurement model fits the data. In EFA, sampling adequacy, measured by the Kaiser-Meyer-Olkin test and sphericity, measured by Bartlett's test, are the technical parameters that are usually reported. The value of the former should be above 0.5 and the statistical significance of the latter should be under 0.05 (Tabachnick and Fidell, 2014). CFA reports several goodness-of-fit statistics. We followed recommendations for reporting multiple fit indices to ensure adequate model fit estimation (Jackson *et al.*, 2009). This includes the ratio of the chi-square value and associated degrees of freedom (χ^2/df); at least one incremental index, such as the comparative fit index (CFI), Tucker-Lewis index (TLI), relative non-centrality index (RNI) or Bollen's delta 2; and at least one residual-based measure, such as the root mean square error of approximation (RMSEA), root mean residual (RMR) or standardized root mean square residual (SRMR). For acceptable fit of the model to the data, it is recommended that the value of χ^2/df be 3 or lower, the values of the incremental indices be above 0.9, and the values of the residual-based measures be below 0.1 (Child, 2006; Ramlall, 2016). The criterion is considered to be fulfilled if all three parameters reach the threshold values, and it is considered to be partially fulfilled if only one or two of the parameters reach these values. Note that these are quite loose thresholds that are not unanimously accepted. However, we aimed to identify the threshold values that are common in the relevant methodological studies.

The quality of reported validity and reliability, as a second category of criteria, pertains to the empirical evidence regarding the measurement instrument that can be assessed based on the statistical parameters reported by the study. These parameters provide insights into the facets of validity and reliability that can be quantitatively evaluated (Table 2).

The evidence for construct validity is partitioned into evidence regarding convergent, discriminant and criterion validity. Convergent validity is usually computed on the basis of factor weights (loadings). In general, higher weights indicate higher convergent validity (DiStefano *et al.*, 2009). The literature usually considers an absolute value of 0.4 for factorial weights or an average variance extracted (AVE) coefficient of higher than 0.5 as critical for convergent validity. Discriminant validity is usually assessed based on the square root of AVE, which should be greater than inter-construct correlations or the correlations between factors, which should be low (Hair *et al.*, 2014) [6].

Criterion validity is another very important – if not the most important – dimension of construct validity that describes the extent to which the measure correlates with an expected outcome or a variable with which it is supposed to be highly correlated (Bollen, 1989; Ferligoj *et al.*, 1995). It is suggested that the correlation be above 0.5 (Fornell and Larcker, 1981). Although criterion validity is one of the strongest pieces of evidence for construct validity, it is usually the most difficult to obtain.

Reliability is another important criterion for the quality of a measurement, and it is usually estimated based on Cronbach's alpha, which should be above 0.7 (Miller, 2009; Nunnally and Bernstein, 1994) or the composite reliability coefficient, which should be above 0.6 (Allen and Yen, 2002). The criterion is considered to be fulfilled if one of the two parameters reaches the threshold value. Social science methodology may include other types of reliability (such as test-retest reliability), but as they are not reported by existing studies (likely because they require more complex, timely and expensive research designs), we did not include them in the table with the criteria.

All the above mentioned criteria can be objectively discerned from published papers. The task of coding was completed by two reviewers, who are experts in social science methodology. They independently read all the papers and extracted the necessary information. In rare cases where the reviewers disagreed on the extracted information, they deliberated until they reached a unanimous decision.

3. Results

3.1 Summary overview of studies

Table 3 shows the basic characteristics of the 19 studies included in the systematic review. An investigation of the geographical characteristics of the studies shows that six studies were conducted in Asia (Saudi Arabia, Malaysia and Oman), five were conducted in Africa (Republic of South Africa and Nigeria), four in North America, two in Europe (Norway and Sweden) and one in Australia. One study is international and was conducted in several countries. All the studies were published after 2007, indicating that the field of measuring and researching information security culture is still quite young, and the majority of the studies were conducted in the past five years. The majority of studies that report the type of organization examined public organizations (e.g. universities and ministries). Finally, there was a large amount of variability of the sample sizes in the studies, ranging from as little as 22 to almost 2,500.

Table 4 provides an overview of the number of items, content of the sample items and unit of analysis of the studies. There were some major differences in the nature of the items between different scales. The items of some scales assume that employees have extensive knowledge of organizational information security and are able to provide answers to relatively difficult questions that pertain to the adherence of management to information security policies (Masrek *et al.*, 2018b; Mokwetli and Zuva, 2018) or assessment of the effectiveness of awareness initiatives (Da Veiga and Eloff, 2010). The items in other scales are straightforward and assess the frequency of activities that affect the information security of an organization, such as sharing passwords and opening emails from unknown senders (Parsons *et al.*, 2014).

Moreover, the items in some scales may be suggestive, asking respondents to assess statements that reveal desirable states, such as intention to protect information technology resources, in an overly direct manner (Rocha Flores and Ekstedt, 2016). Several items that pertain to adherence to information security policies simply ask respondents to assess their adherence to such policies without testing the assumption that respondents know the content of organizational security policies.

Other scales contain items with rather vague statements regarding the relationship between the organization and individual employees, such as "my company makes employees aware of its security policies and regulations" (D'Arcy and Greene, 2014). Although this statement reflects an important element of information security culture, it may be based on incorrect assumptions, and thus, may not produce valid answers.

Our overview of example items indicates that approaches to measure information security culture vary significantly. Although face validity is a subjective assessment, some items may be very difficult for ordinary employees to answer and others may be prone to the ecological fallacy (Ess and Sudweeks, 2001) or social desirability bias (Hays *et al.*, 1989), which are not uncommon in the early stages of developing metrics in a nascent research field.

Our investigation also reveals that there are two ontologically different ways of understanding the phenomenon of information security culture. As shown in Table 4, one set of studies (11 studies) understands it as a higher-level construct comprising

Name of scale/study[a]	Authors/year of publication	Country[b]	Type of organization	Sample size
Security culture	Alharbi *et al.* (2017)	Kingdom of Saudi Arabia (KSA)	NA	$n = 625$
Information security culture framework (ISCF)	Al Hogail (2015)	Kingdom of Saudi Arabia (KSA)	Financial government organization, small to medium-sized enterprise and multi-national corporation	$n = 22–52$[c]
Information security culture	AlKalbani *et al.* (2015)	Oman	Public organizations offering e-government services	$n = 294$
Information security culture	Alnatheer *et al.* (2012)	Kingdom of Saudi Arabia (KSA)	200 organizations in Saudi Arabia (sample covers all regions, as well as types and sizes of organizations)	$n = 254$
Security culture	Brady (2011)	USA	Academic medical centers	$n = 76$
Security culture[a]	Chen *et al.* (2015)	USA	Four companies in Midwestern USA	$n = 100$
Information security culture	Choi (2019)	South Korea	Korean companies	$n = 305$
Information security culture framework (ISCF)	Da Veiga and Eloff (2010)	Republic of South Africa (RSA)	Audit and advisory assignments	$n = 1,085$
Information security culture assessment (ISCA)	Da Veiga and Martins (2015a, 2015b)	Republic of South Africa (RSA)	International financial organization (across 12 countries)	$n = 1,571$-2,320[c] (2006, 2008, 2010 and 2013)
Information protection culture assessment (IPCA)	Da Veiga and Martins (2015a, 2015b)	Republic of South Africa (RSA)	International financial organization (across 12 countries)	$n = 2,159$, 2,320[c] (2010 and 2013)
Security culture	D'Arcy and Greene (2014)	USA	Various organizations throughout the Mid-Atlantic region	$n = 127$
Organizational security culture	Khan and Alshare (2019)	USA	Midwestern University	$n = 195$
Security culture	Knapp *et al.* (2007)	Different countries in the world	Various companies from different industries	$n = 740$
Collection security culture	Maidabino and Zainab (2012)	Nigeria	Four Nigerian university libraries	$n = 61$
Information security culture	Masrek *et al.* (2018a, 2018b)	Malaysia	Technology department of Malaysian federal ministries	$n = 292$
ICT security culture	Mokwetli and Zuva (2018)	Republic of South Africa (RSA)	Small, medium and micro enterprises (SMME) in Gauteng	$n = 647$
Organizational security culture scale	Nævestad *et al.* (2018)	Norway	NA	$n = 323$, 446[c] (2014 and 2016)

Table 3.
Basic information of the selected studies

(*continued*)

Name of scale/study[a]	Authors/year of publication	Country[b]	Type of organization	Sample size
Organizational information security culture	Parsons *et al.* (2014)	Australia	NA	$n = 500$
Information security culture	Rocha Flores and Ekstedt (2016)	Sweden	NA	$n = 1{,}583$

Notes: [a]Where the scales are not explicitly named, the name of the study is provided. [b]The country refers to the location where the study was conducted. [c]Sample size data refers to different organizations or measurements at different time periods. Not available (NA)

Table 3.

several dimensions. In this case, information security culture is not measured directly, but via a set of dimensions that pertain to different parts of information security culture. In such cases, each dimension is represented by a set of measurement items. In contrast, the other set of eight studies measures information security culture directly via a set of measurement items. In other words, these eight studies do not consider dimensions of information security culture, but measure it as a unidimensional concept. The evaluation framework is the same; the only difference is that the "dimensions" column in Table 4 is empty for these studies.

3.2 Evaluation results

We presented the results regarding the rigor of operationalization and reported empirical evidence for the quality of measurement separately. In addition, we divided the studies into two sub-groups for evaluation according to whether they measure information security culture as a multidimensional construct or unidimensional concept. Note that when measuring information security as a unidimensional concept, the total number of operationalization criteria is 11 instead of 12 because the criterion "presence of definitions of dimensions" is irrelevant in these cases.

Figure 2 reports the results of the evaluation of evidence for the rigor of operationalization and demonstrates that none of the studies fulfill all 12 criteria. Also, no individual criterion is fulfilled by all 19 studies. Only three studies do not provide an essential definition of the concept, which is a crucial starting point for operationalization and a baseline for determining face and content validity. The majority of studies report the source of items, while five do not. Approximately half of the studies report conducting expert review and pilot testing. Only four studies comprehensively report the conceptual background of the scale and conduct both expert review and pilot testing.

The documentation regarding other criteria of the rigor of operationalization, such as the presence of sample characteristics, basic descriptive statistics and a correlation matrix, is quite limited. Also, data on the suitability of EFA and the fit of the measurement model in CFA are scarce. For instance, model fit parameters are reported by only seven studies, and of those, three do not report all fit parameters, as suggested by the literature.

The scale of information security culture developed by Alnatheer *et al.* (2012) obtained the highest score for the rigor of operationalization of information security culture as a multidimensional construct, followed by the information security culture scale (Masrek *et al.*, 2018b) and the information security culture assessment (ISCA) (Da Veiga and Martins, 2015a). The common feature of these studies is that they

Name of scale/study	Dimensions	Total no. of items	Example of item	Unit of analysis
Security culture (Knapp *et al.*, 2007)[a]		6	"A culture exists that promotes food security practices"	Constituent
ISCF (Da Veiga and Eloff, 2010)	Leadership and governance, security management and operations, security policies, security program management, user security management, technology protection and operations, change	85	"I believe it is necessary to protect information to achieve the business strategy" *(leadership and governance)*	Employee
Security culture (Brady, 2011)[a]		10	NA	Health care professionals
Information security culture (Alnatheer *et al.*, 2012)	Top management (involvement), information security policy, information security training, information security awareness, information security ownerships	19	"I am aware of my information security roles and responsibilities" *(information security awareness)*	Employee
Collection security culture (Maidabino and Zainab, 2012)	Aware of accidents, perceptions, values and attitudes toward collection protection, awareness of obstacles	42	"Staff knows their individual library security responsibility" *(perceptions, value and attitudes)*	Employee
Security culture (D'Arcy and Greene, 2014)	Top management commitment, security communication, computer monitoring	12	"I believe that my organization reviews logs of employees' computing activities on a regular basis" *(computer monitoring)*	Computer-using professionals (without background in information security)
ISCF (Al Hogail, 2015)	Strategy, technology, people, organization, environment	72	"The elements of information security strategies clearly state what is expected from me" *(strategy)*	Employee
Information security culture (AlKalbani *et al.*, 2015)	Management commitment, accountability, information security awareness	17	NA	Employee
Security culture	Security policy, security education, training and	17	"The overall environment in my	Employee

(*continued*)

Table 4.
Overview of scales and examples of items

Name of scale/study	Dimensions	Total no. of items	Example of item	Unit of analysis
(Chen *et al.*, 2015)[a]	awareness (SETA) programs, security monitoring, security culture		organization fosters security-minded thinking in all our actions" (*security culture*)	
ISCA (Da Veiga and Martins, 2015a)	Information asset management, information security policies, change management, trust, information security management, information security program, user management, information security leadership, training and awareness	45	"I believe I am responsible for the protection of company's information assets (e.g. information and computer resources)" *(user management)*	Employee
IPCA (Da Veiga and Martins, 2015b)	Information security commitment, management buy-in, information security necessity and importance, information security policy effectiveness, information security accountability, information usage perception	55	"I believe the information security awareness initiatives are effective" *(information security commitment)*	Employee
Organizational information security culture (Parsons *et al.*, 2014)[a]		NA	NA	Employee (using a computer or portable device at their work; organization with a formulated information security policy)
Information security culture (Rocha Flores and Ekstedt, 2016)[a]		7	"My colleagues would warn me if they saw me taking risks" *(information security culture)*	Employee
Security culture (Alharbi *et al.*, 2017)[a]		2	The government considers information security an important priority	Citizen
Information security culture	Management support, policy and procedure, compliance, awareness, budget, technology	48	NA	IT employee

(*continued*)

Table 4.

Name of scale/study	Dimensions	Total no. of items	Example of item	Unit of analysis
(Masrek *et al.*, 2018b)				
ICT security culture (Mokwetli and Zuva, 2018)	Information security commitment, information security importance, information security policy effectiveness, information security directives, information security monitoring compliance, information security consequences, information security responsibility, information security training, information security budget practice, information security investment, information technology capability, information technology compatibility	21	NA	Business owners/ managers and their direct level reporting structures within the SMMEs environment
Organizational security culture scale (Nævestad *et al.*, 2018)	Management commitment, employees' attitudes, reporting culture and reactions to incident reporting, safety training and education, general information security issues	25	Everyone is informed of any changes that may impact information security *(safety training and education)*	Employee
Information security culture (Choi, 2019)[a]		7	NA	Employee
Organizational security culture (Khan and Alshare, 2019)[a]		2	Information security is a key norm shared by the members in our organization	Employee

Table 4.

Notes: [a]Security culture is measured as one-dimensional concept. Not available (NA)

provide a conceptual background for the development of items and report descriptive statistics and inter-item or factor correlations. Among the scales that measure information security culture as a unidimensional concept, the highest score was obtained by the security culture scale developed by Chen *et al.* (2015). They provide a definition of the concept, clearly document the process of producing items and provide sample characteristics and correlation statistics.

	Essential definition of the concept	Definition of the components of the concept	Source of items	Expert review	Pilot test of items	Sample size	Response rate	Sample characteristics	Descriptive statistics of the items or factors	Correlations between factors/items	EFA	CFA	Total fulfilled criteria	% fulfilled criteria
Information security culture (Alnatheer et al., 2012)	●	●	●	●	●	●	●		●	●	●	●	11	91.7
Information security culture (Masrek et al., 2018b)	●	●	●		●	●	●	●	●	●			9	75.0
ISCA[2] (Da Veiga and Martins, 2015a)	●	●	●			●	◐		●	●	●	◐	8	66.7
ISCF (Da Veiga and Eloff, 2010)	●	●	●	●	●	●	●					◐	7.5	62.5
Organisational security culture scale (Nævestad et al., 2018)	●		●		●	●	●		●		●		7	58.3
ISCF (AlHogail, 2015)	●	●	●	●		○	◐		●	●		◐	7	58.3
IPCA[2] (Da Veiga and Martins, 2015b)	●	●	●			●	◐		●		?		5.5	45.8
Security culture (D'Arcy and Greene, 2014)	●		●			●	●	◐		●			5.5	45.8
Information security culture (AlKalbani et al., 2015)	●	●		●		●	○					●	5	41.7
ICT security culture (Mokwetli and Zuva, 2018)	●					●		◐	●		●		4.5	37.5
Collection security culture (Maidabino and Zainab, 2012)	●				●	○	●		●		◐		4.5	37.5
Security culture (Chen et al., 2015)[1]	●		●	●	●	●		●		●		●	8	72.7
Information security culture (Rocha Flores and Ekstedt, 2016)[1]	●		●	●	●	●	●	◐		●			7.5	68.2
Organisational information security culture (Parsons et al., 2014)[1,2]			●	●	●	●		◐	●	●			6.5	59.1
Organisational security culture (Khan and Alshare, 2019)[1]	●					●		●	●	●			5	45.5
Security culture (Alharbi et al., 2017)[1]	●		●			●		●		●			5	45.5
Security culture (Knapp et al., 2007)[1]			●	●	●	●	○					●	5	45.5
Information security culture (Choi, 2019)[1]			●			●		●		●			4	36.4
Security culture (Brady, 2011)[1]	●					●	○			●			3	27.3

Notes: [1] Security culture is measured as (an independent) unidimensional concept;
[2] Information from several articles is collected into one evaluation (IPCA: Da Veiga and Martins, 2015b; Da Veiga, 2016 and Da Veiga, 2018; Information security culture: Masrek *et al*., 2018a and Masrek *et al*., 2018b; ISCA: Da Veiga and Martins, 2015a and Martins and Da Veiga, 2015; Organisational information security culture: Parsons et al., 2014 and Parsons *et al*., 2015)
Legend:
●: the criterion is fulfilled
◐ : the criterion is partially fulfilled
?: limited evidence because of no numerical data and only an interpretation
Empty cell: data is not available

Figure 2.
Evaluation of rigor of operationalization

The results of the examination of empirical evidence for the validity and reliability of the studied scales are presented in Figure 3. Similar to above, none of the studies fulfill all the statistical criteria and none of the criteria are fulfilled by all 19 studies. Reliability is a commonly reported criterion of measurement quality; 14 of 18 scales fulfill the empirical condition as internal consistency and composite reliability are higher than 0.7 and 0.6, respectively. This criterion is only partially met by ICT security culture scale (Mokwetli and Zuva, 2018), the information security culture framework scale (Al Hogail, 2015) and the ISCA scale (Da Veiga and Martins, 2015a). The security culture scale developed by Alharbi *et al.* (2017) does not meet the thresholds. Furthermore, six studies do not mention or test any

	Convergent validity	Discriminant validity	Criterion validity	Reliability	Total fulfilled criteria	% fulfilled criteria
Information security culture (Masrek et al., 2018b)[3]	●	●		●	3	75.0
Information security culture (AlKalbani et al., 2015)	●	●		●	3	75.0
Security culture (D'Arcy and Greene, 2014)	●	●		●	3	75.0
Information security culture (Alnatheer et al., 2012)	●	●		●	3	75.0
Organisational security culture scale (Nævestad et al., 2018)	●			●	2	50.0
IPCA (Da Veiga and Martins, 2015b)[3]				●	1	25.0
Collection security culture (Maidabino and Zainab, 2012)				●	1	25.0
ISCF (Da Veiga and Eloff, 2010)				●	1	25.0
ICT security culture (Mokwetli and Zuva, 2018)				◐	0.5	12.5
ISCF (Al Hogail, 2015)				◐	0.5	12.5
ISCA (Da Veiga and Martins, 2015a)[3]	○			◐	0.5	12.5
Organisational security culture (Khan and Alshare, 2019)[2]	●	●		●	3	75.0
Information security culture (Rocha Flores and Ekstedt, 2016)[2]	●	●		●	3	75.0
Security culture (Chen et al., 2015)[2]	●	●		●	3	75.0
Security culture (Knapp et al., 2007)[2]	●	●		●	3	75.0
Information security culture (Choi, 2019)[2]	●	●			2	50.0
Security culture (Alharbi et al., 2017)[2]	●	●		○	2	50.0
Organisational information security culture (Parsons et al., 2014)[2,3]				●	1	25.0
Security culture (Brady, 2011)[2]		○		●	1	25.0

Notes: [1]In the case of a partially fulfilled criterion, half of the point was taken into account in the total score; [2]Security culture is measured as (an independent) unidimensional concept; [3]Information about characteristics of the same scale is collected from several articles (Information security culture: Masrek *et al*., 2018a and Masrek *et al*., 2018b; ISCA: Da Veiga and Martins, 2015a and Martins and Da Veiga, 2015; IPCA: Da Veiga and Martins, 2015b; Da Veiga, 2016 and Da Veiga, 2018; Organisational information security culture: Parsons et al., 2014 and Parsons *et al*., 2015)
Legend:
●: the criterion is fulfilled
○: the criterion is not fulfilled because it does not reach threshold values
◐ : the criterion is partially fulfilled
Empty cell: information is not available

Figure 3.
Evaluation of reported validity and reliability

aspect of construct validity and criterion validity (concurrent or predictive) is absent from all the studies. Two scales do not meet the thresholds for convergent and discriminant validity.

Four scales fulfil three out of four criteria for reported quality of measurement when information security culture is considered a multidimensional construct (Masrek *et al.*, 2018a, 2018b; AlKalbani *et al.*, 2015; D'Arcy and Greene, 2014; Alnatheer *et al.*, 2012). Regarding the scales that measure information security culture as a unidimensional concept, four studies – including the oldest in this analysis – received identical evaluation scores (Khan and Alshare, 2019; Rocha Flores and Ekstedt, 2016; Chen *et al.*, 2015; Knapp *et al.*, 2007).

3.3 Associations of evaluation scores with type of scale, year of publication and journal impact

In this section, we provide an aggregated perspective on the studies that measure information security culture and investigate whether the rigor of operationalization and reported validity and reliability of measurement instruments are associated with the type of scale, year of publication and journal impact. Table 5 reports the percentages of criteria that are fulfilled by each category of articles.

Studies that measure information security culture with a multidimensional scale fulfill, on average, 56.1% of the operationalization criteria, while studies that measure information security culture as a unidimensional concept fulfill 45.8% of these criteria. The empirical evidence for the validity and reliability of multidimensional scales is not as strong, fulfilling 42.0% of criteria, compared to studies with a unidimensional construct, which fulfilled 56.3% of criteria.

The results show that the reported validity and reliability somewhat improve with the recency of studies, but this is not the case for the rigor of operationalization. In general, evidence for the rigor of operationalization seems to be stronger than that for empirical reliability and validity. Regarding the rigor of operationalization, the highest score was

	No. of studies	% fulfilled criteria (operationalization)	% fulfilled criteria (validity and reliability)
Type of scale			
Information security culture as a multidimensional scale	11	56.1	42.0
Information security culture as a unidimensional scale	8	45.8	56.3
Year of publication			
up to 2013	5	52.6	45.0
2014–2015	7	55.5	42.9
2016–2017	2	56.8	62.5
2018–2019	5	50.9	52.5
Impact factor of journal			
N.A	5	51.7	47.5
<0.5	7	49.4	57.1
0.5 < IF < 1.0	4	61.7	27.5
1.0 < IF < 2.0	2	56.5	31.3
IF > 2.0	1	68.2	75.0

Table 5. Associations between evaluation scores and type of scale, year of publication and impact factor of journal

obtained by studies from 2016–2017 (56.8%). Interestingly, studies that were published before 2014 fulfill, on average, 52.6% of the criteria, while studies that were published in the past two years fulfilled 50.9% of the criteria. Regarding reported reliability and validity, the highest score was obtained by studies from 2016–2017 (62.5%), while the lowest score was obtained by studies from 2014–2015 (42.9%).

The association between journal impact and evidence for validation of the scale seems to be stronger. For studies published in journals with the highest impact factor, an average of 68.2% of the rigor of operationalization criteria and 75.0% of the validity and reliability criteria were fulfilled. In contrast, for studies published in journals with the lowest impact factor, the evaluated scales fulfilled only 49.4% of the rigor of operationalization criteria and 57.1% of the validity and reliability criteria. Interestingly, the evidence for validity and reliability in studies from journals with lower impact factors is better than the evidence in studies from journals with an impact factor of 0.5–2. In publications without an impact factor (mostly papers published as part of conference proceedings), 51.7% of the rigor of operationalization criteria and 47.5% of the validity and reliability criteria were fulfilled.

4. Discussion

The main aims of our research were to identify the studies that propose survey-based scales of information security culture, to provide an overview of these studies and to examine the reported evidence regarding the rigor of operationalization and the quality of the scales' measurement in terms of empirical support for validity and reliability. During the selection process, we discovered 11 studies that measure information security culture as a multidimensional construct and 8 studies that measure it as a unidimensional concept. The results of our systematic review show that the evidence for validation of the scales is somewhat limited. Validity is one of the most important aspects of a high-quality measurement instrument, and providing an essential definition is a necessary, but by no means sufficient, condition of validity. The majority of scales provide definitions of the concept/construct, which is immensely important for further development of the scales and the research field in general. Moreover, most of the studies provide an underlying logic for the development of items, which allows other researchers to make subjective assessments of the face and content validity of the proposed scales. In contrast, the reported statistical tests of the scales are rather limited for various factor-analytical procedures, and empirical insights into validity are more rare or even completely absent in the case of criterion validity.

4.1 A need for predictive validity

The field of research and measurement of information security culture is relatively young; thus, it is expected that studies are not (yet) investigating criterion validity because theory does not offer much background for expected correlations. Nevertheless, this type of validity is one of the most important to empirically prove that the proposed scale is measuring what it is supposed to (Carmines and Zeller, 1979; DeVellis, 2016). This is even more important when the outcomes of measurements are taken as a baseline for decision-making. The need to establish criterion validity is particularly important in light of the (rather subjective) observation that face and content validity may be problematic in cases with several scales. As suggested in Section 3 of this paper, some items seem to explicitly relate to desirable states and may invite social desirability bias. For example, in one study, the item "The information security policy is understandable by all employees irrespective of their ranks" received a mean agreement of 4.7 (on a scale of 1 to 5), and the average of all other items was higher than 4.3. This may indicate an excellent information security culture, but the validity

of such a result may be questionable. If the majority of employees understand and comply with information security policies, then the human factor would not be such a critical issue in many contemporary organizations (Tsohou *et al.*, 2015). Thus, it is crucial to provide evidence that responses to items on the information security culture scale correlate with employee actual behavior. In other words, if the measures of information security culture indicate a strong information security culture, then the employees should be very unlikely to be the cause of security incidents. If such a correlation is absent, then the items are likely subject to social desirability bias or another type of bias. As surveys are often conducted in organizational contexts, respondents may feel that they are being observed, and thus, provide responses that are more in line with expectations than with reality. Thus, the most problematic aspect of seminal information security scales is their sensitivity to social desirability bias (Lebek *et al.*, 2014).

Empirical tests based on factor analysis and other statistical methods that provide insights into reliability and convergent and discriminant validity are unable to solve this issue. Moreover, although these tests can provide evidence of high empirical validity and reliability, such results can be subject to different biases, and thus, be invalid. Further studies should, therefore, aim to demonstrate whether information security culture is indeed correlated with behavioral or organizational outcomes in terms of information security and organizational security.

4.2 Information security culture as an umbrella concept

One of the unintended consequences of our research is our demonstration of the complexity and inclusivity of the concept of information security culture. The majority of the analyzed scales are multidimensional, suggesting that the concept of information security culture is comprising several (from 3 to more than 10) dimensions that pertain to various aspects of information security culture in organizations. The lack of documentation and satisfaction of convergent and discriminant validity, however, suggests that researchers need to invest more energy into defining essential dimensions of information security culture, making a clearer distinction between them, and establishing theoretical relations between them. The studies that were identified in this systematic review base their conceptual apparatus on various theoretical backgrounds, and because this is a relatively young field of study, it is not surprising that the concept of information security culture has not yet been adopted by all researchers focusing on the human factor. To further establish and recognize the importance of researching the human factor in information security, developing a common concept is necessary. Information security culture offers a valid baseline, but higher agreement on the definition of this concept is needed. Such a common conceptual background is important to develop and optimize valid items, not only for survey-based measurements of information security culture but also for measurement based on automated data collection procedures.

4.3 Practical implications

The results of our systematic review have several practical implications for the empirical research on information security culture and for organizations aiming to assess and measure the strength of the "human firewall." We have identified several scales with comprehensive documentation and good measurement characteristics, but some important aspects of validity and reliability have not yet been established. Consequently, it is not advisable for researchers and practitioners to simply replicate these scales, as it is not yet clear, which scales correlate with the actual security of organizations. This raises a question: how should a researcher decide, which scale to use? First, we suggest that studies should build upon a synthesis of existing scales,

particularly those that demonstrate strong evidence of validation. At the same time, this research field is relatively young, which means that additional efforts must be made in the pre-testing phase. This pre-testing should not only include a pilot study of the proposed instrument but also the application of a mixed-method design, in which qualitative insight into the meanings respondents ascribe to certain items is essential. Ideally, a scale should be tested with a multi-trait, multi-method approach (Saris, 1995), which is time- and cost-intensive but provides the most comprehensive empirical insights into different facets of validity and reliability.

Scales of information security culture often have numerous measurement items, which can be a drawback because a larger number of items impose a burden on respondents and can lead to lower response rates (Callegaro *et al.*, 2015; Rubin, 2004). Therefore, it is important for researchers to consider two implications. First, the pre-testing phase should aim to reduce the number of items per dimension to an essential minimum (at least three items per dimension allows for an advanced statistical test of different types of validity and reliability). Second, not all dimensions of information security culture may be relevant for specific types of industries or organizations, and thus, the measurement phase could be limited to certain dimensions. Moreover, in such cases, it might be advisable for researchers to apply unidimensional scales of information security culture. Although these scales are in the minority, our study demonstrated that their quality is similar to that of multidimensional scales.

Our study has some direct and indirect implications for organizations that aim to measure information security culture among their employees. First, an overview of the scales showed that some, especially multidimensional scales of information security culture, tend to include many aspects of "the human factor" in organizational information security. This means that organizations can measure not only the knowledge and behaviors of employees regarding security but also their values, beliefs and attitudes toward security. With this information, organizations can enhance their security awareness programs, which usually address only employees' knowledge, with a broader and more tailored curriculum. If, for example, regular measurements detect that attitudes toward security are becoming more negative while knowledge of security remains the same, organizations can take this as a clear message that the information security culture program should primarily address the mechanisms for improving attitudes toward security. For this purpose, multidimensional scales of information security culture are especially relevant, as they allow for detailed insight into organizational security culture (under the condition that they are valid and reliable). The drawback, however, is that these scales usually contain a large number of items and impose more burden on employees. Organizations thus, need to find a good balance between an informative number of items and the cost of conducting surveys. The testing phase is of utmost importance to determine the parsimonious number of items, and thus, minimize employee burden without impacting the quality of measuring.

This sort of balancing also applies to questions about how often the information security culture should be measured within organizations. The metrics obtained by a single measure are informative for decision-makers in the organization, but the highest value comes from repeating the measures with certain time intervals. Changes and trends in information security culture (especially in its dimensions) provide important metrics for tailoring information security culture programs and monitoring the impact of security programs. However, regular solicitation of surveys and recruiting of employees requires careful implementation and an understanding of costs. Solicitation should be done in such a way that employees do not feel that their answers will have any sort of impact on their status in the organization. If this condition is not met, the answers will likely be subject to social

desirability bias. Therefore, it is advisable to pre-test items for social desirability bias by correlating them with the social desirability scale (Hays *et al.*, 1989).

An important indirect implication of our study is that the scales of information security culture seem to be relevant for the majority of industries. As the unit of analysis is typically an "ordinary" employee, the content of items seems independent of the industry, and thus, is applicable to any kind of organization. However, this means that all employees in the organization need to be solicited for the survey instrument, not only a sample of them. This is an issue of implementation that requires a balance to be found between the burden on the employees and on the organization. However, the results can provide crucial information for decision-makers to take adequate action. For example, an inter-unit analysis might reveal the emergence of specific security sub-cultures, which could pose great risks to organizational security (Da Veiga and Martins, 2017).

4.4 Limitations

Our systematic review is subject to certain limitations, which warrant further research. First, as explained above, our search strategy was limited to a set of terms that, in our opinion, returned a comprehensive set of studies that measure information security culture. However, it is possible that there are studies that measure information security culture but do not use this term or related wording in the title or abstract of the study. Second, the reader should be aware that the low evaluation scores of some scales do not necessarily imply that such scales are of low quality. This may be a consequence of the fact that some authors do not report some psychometric characteristics of the scales or other elements of the operationalization process. Our evaluation gave a score of 0 both when a criterion is not fulfilled and when there was no reported data that could be used to check the fulfillment of the criterion. This was especially common for convergent and discriminant validity, as authors rarely report the results that would allow for the evaluation of these two validities. Third, our evaluation was limited to elements related to the rigor of validation, which can be quantified and objectively coded. We did not inspect an important perspective of validation that pertains to the (theoretical) issues of internal and external validity. In our evaluation, we followed the suggested methodological standards, which instruct authors to supply publications of new or adapted scales with all the necessary information. However, these standards were developed to fit any kind of measurement scale in the social sciences, from those used in organizational studies to those used in psychology and health studies. In the field of information security culture, there is currently no consensus on the "gold standards" of measurements, specific to this field. In the future, we should, however, strive for the development of such a common framework that would, for instance, dedicate special attention to testing of the social desirability effect during the validation of scales. This issue seems to be among the critical ones in the field of measuring information security culture (Lebek *et al.*, 2014). Finally, the methodological guidelines regarding threshold values and some of the statistical parameters are not universally accepted. Consequently, the values are always somewhat arbitrary, although we tried to use a common denominator among different studies.

5. Conclusion

By answering the first research question, we managed to identify 19 studies that introduce a measure of information security culture, either as a multidimensional scale (11 studies) or a unidimensional one (8 studies). Analysis of the second research question reveals that the majority of scales achieve modest rigor of operationalization and none of the studies fulfills all the criteria of our evaluation framework. Similarly, the analysis of the third research

question shows that the reported validities and reliabilities of the identified scales are somewhat modest. Especially problematic seems to be the absence of testing and reporting the criterion validity.

The measurement of information security culture thus, still has room for improvement, as none of the studies fulfilled all the methodological criteria used in this systematic review. Moreover, several studies do not report some essential elements of the operationalization process that would enable replicability. The lack of convincing evidence for validation of information security culture scales is not surprising because the field is relatively young and interest in the concept of information security culture has significantly increased only in the past few years after scholars realized that the concept of information security awareness was too narrow to address the complexity of the human factor in information security (Metalidou *et al.*, 2014). However, it is necessary for optimization and testing of scales in various organizational contexts to continue because the research field needs to reach at least some agreement on the basic set of scales to address information security culture and guarantee comparison among studies, organizations and various other contexts. Some degree of consolidation of the field, the scales used and the standards for measurement in the field are urgently needed. In addition, organizations need valid scales to assess the true state of their information security culture and act accordingly. This study represents a step in these directions.

Notes

1. Social science methodology often uses the term "scale" to refer to a measurement instrument, which is a collection of items that are (on an analytical level) combined into a composite score and intended to assess latent concepts that are not readily observable by direct means (DeVellis, 2016).
2. The term "cybersecurity culture" was included in the search because some authors relate it to organizational security culture, although cybersecurity culture is usually understood as a broader concept.
3. Students and employees have free access to the database of scientific literature because of their relationship to the University of Ljubljana. "Open access" refers to online research results that are freely available in different databases without copyright or licensing restrictions.
4. Note that in four cases (Da Veiga and Martins, 2015a, 2015b; Masrek *et al.*, 2018b; Parsons *et al.*, 2014), we decided to aggregate the information about the same measurement instrument from multiple studies into one evaluation. This was done in cases where more than one study by the same author (s) reports measurement characteristics for identical samples. For example, information about Da Veiga's instruments regarding information security culture was reported in several articles that rely on the same data. In such cases, information about a measurement instrument was not exhaustively documented in a particular article; rather, different aspects were presented in various articles. Because those scales were tested using the same data (same sample size and demographic data), we decided to summarize the criteria into one evaluation.
5. Here, we limit our analysis to aspects that are quantified and can be objectively coded, such as model fit parameters and we do not code the quality of qualitative information, such as the quality of definitions.
6. In cases that use a multidimensional concept of information security culture, we estimate factorial weights, model fit and convergent and criterion validity for the entire instrument. For cases that use a unidimensional concept, we consider the statistics reported for only that dimension (if applicable).

References

Al Hogail, A. (2015), "Cultivating and assessing an organizational information security culture; an empirical study", *International Journal of Security and Its Applications*, Vol. 9 No. 7, pp. 163-178.

Al Hogail, A. and Mirza, M. (2015), "Organizational information security culture assessment", paper presented at The 2015 International Conference on Security and Management (SAM'15), 27-30 July, Las Vegas, available at: http://worldcomp-proceedings.com/proc/p2015/SAM_contents.html (accessed 15 June 2019).

Alharbi, N., Papadaki, M. and Dowland, P. (2017), "The impact of security and its antecedents in behaviour intention of using e-government services", *Behaviour and Information Technology*, Vol. 36 No. 6, pp. 620-636.

AlKalbani, A., Deng, H. and Kam, B. (2015), "Organisational security culture and information security compliance for E-Government development: the moderating effect of social pressure", paper presented at The Pacific Asia Conference on Information Systems (PACIS), 5-9 July, Singapore, available at: https://pdfs.semanticscholar.org/2892/fe0931830eb5665e5b1614440d965978926f.pdf?_ga=2.6402546.1915429506.1576068243-1272771706.1576068243 (accessed 7 July 2019).

Allen, M.J. and Yen, W.M. (2002), "Introduction to measurement theory", available at: http://books.google.si/books?id=MNUpY_csc6cC (accessed 18 June 2019).

Alnatheer, M., Chan, T. and Nelson, K. (2012), "Understanding and measuring information security culture", paper presented at The Pacific Asia Conference on Information Systems (PACIS), 11-15 July, Ho Chi Minh City, Vietnam, available at: www.pacis-net.org/file/2012/PACIS2012-005.pdf (accessed 29 June 2019).

American Educational Research Association – AERA, American Psychological Association – APA, National Council on Measurement in Education – NCME (1992), *Standards for Educational and Psychological Testing*, APA, Washington, DC.

American Psychological Association (APA) (2020), *Publication Manual of the American Psychological Association*, APA, Washington, DC.

Bollen, K.A. (1989), "A new incremental fit index for general structural equation models", *Sociological Methods and Research*, Vol. 17 No. 3, pp. 303-316.

Brady, J.W. (2011), "Securing health care: assessing factors that affect HIPAA security compliance in academic medical centers", in *Proceedings of the 2011 44th HI International Conference on System Sciences*, IEEE Computer Society, Washington, DC, pp. 1-10.

Budge, J., O'Malley, C., McClean, C., Barnes, M., Sebastin, S. and Nagel, B. (2018), *Instill a Security Culture by Elevating Communication*, Forrester Research, Cambridge, MA.

Callegaro, M., Lozar Manfreda, K. and Vehovar, V. (2015), *Web Survey Methodology*, SAGE, Thousand Oaks.

Carmines, E.G. and Zeller, R.A. (1979), *Reliability and Validity Assessment*, SAGE, Beverly Hills.

Chen, Y.A.N., Ramamurthy, K.R.A.M. and Wen, K.W. (2015), "Impacts of comprehensive information security programs on information security culture", *Journal of Computer Information Systems*, Vol. 55 No. 3, pp. 11-19.

Child, D. (2006), *The Essentials of Factor Analysis*, A&C Black, London.

Choi, Y. (2019), "Workplace violence and social engineering among Korean employees", *International Journal of Asian Business and Information Management*, Vol. 10 No. 1, pp. 26-37.

Da Veiga, A. (2016), "A cybersecurity culture research philosophy and approach to develop a valid and reliable measuring instrument", *In Science and Information (SAI) Computer Conference*, Curran Associates, New York, NY, pp. 1006-1015.

Da Veiga, A. (2018), "An approach to information security culture change combining ADKAR and the ISCA questionnaire to aid transition to the desired culture", *Information and Computer Security*, Vol. 26 No. 5, pp. 584-612.

Da Veiga, A. and Eloff, J.H.P. (2010), "A framework and assessment instrument for information security culture", *Computers and Security*, Vol. 29 No. 2, pp. 196-207.

Da Veiga, A. and Martins, N. (2015a), "Improving the information security culture through monitoring and implementation actions illustrated through a case study", *Computers and Security*, Vol. 49, pp. 162-176.

Da Veiga, A. and Martins, N. (2015b), "Information security culture and information protection culture: a validated assessment instrument", *Computer Law and Security Review*, Vol. 31 No. 2, pp. 243-256.

Da Veiga, A. and Martins, N. (2017), "Defining and identifying dominant information security cultures and subcultures", *Computers and Security*, Vol. 70, pp. 72-94.

D'Arcy, J. and Greene, G. (2014), "Security culture and the employment relationship as drivers of employees' security compliance", *Information Management and Computer Security*, Vol. 22 No. 5, pp. 474-489.

DeVellis, R.F. (2016), *Scale Development: Theory and Applications*, SAGE, Newbury Park.

Dillman, D.A., Smyth, J.D. and Christian, L.M. (2014), *Internet, Phone, Mail, and Mixed-Mode Surveys: The Tailored Design Method*, John Wiley and Sons, NJ.

DiStefano, C., Zhu, M. and Mindrila, D. (2009), "Understanding and using factor scores: considerations for the applied researcher, practical assessment", *Research and Evaluation*, Vol. 14 No. 20, pp. 1-11.

Ess, C. and Sudweeks, F. (2001), *Culture, Technology, Communication: Towards an Intercultural Global Village*, Suny Press, New York, NY.

European Union Agency for Network and Information Security (ENISA) (2017), "Cyber security culture in organisations", available at: www.enisa.europa.eu/publications/cyber-security-culture-in-organisations (accessed 15 July 2019).

Ferligoj, A., Leskošek, K. and Kogovšek, T. (1995), *Zanesljivost in Veljavnost Merjenja*, Fakulteta za družbene vede, Ljubljana.

Fornell, C. and Larcker, D.F. (1981), "Evaluating structural equation models with unobservable variables and measurement error", *Journal of Marketing Research*, Vol. 18 No. 1, pp. 39-50.

Furnell, S. and Thomson, K.L. (2009), "From culture to disobedience: recognising the varying user acceptance of IT security", *Computer Fraud and Security*, No. 2, pp. 5-10.

Gordon, L.A. and Loeb, M.P. (2005), *Managing Cybersecurity Resources: A Cost-Benefit Analysis*, McGraw-Hill, New York, NY.

Hair, J.F., Black, W.C., Babin, B.J. and Anderson, R.E. (2014), *Multivariate Data Analysis*, (7th ed.), Pearson, Harlow.

Hays, R.D., Hayashi, T. and Stewart, A.L. (1989), "A five-item measure of socially desirable response set", *Educational and Psychological Measurement*, Vol. 49 No. 3, pp. 629-636.

Hinkin, T.R. (1998), "A brief tutorial on the development of measures for use in survey questionnaires", *Organizational Research Methods*, Vol. 1 No. 1, pp. 104-121.

Jackson, D.L., Gillaspy, J.A., Jr,. and Purc-Stephenson, R. (2009), "Reporting practices in confirmatory factor analysis: an overview and some recommendations", *Psychological Methods*, Vol. 14 No. 1, pp. 6-23.

Khan, H.U. and AlShare, K.A. (2019), "Violators versus non-violators of information security measures in organizations – a study of distinguishing factors", *Journal of Organizational Computing and Electronic Commerce*, Vol. 29 No. 1, pp. 4-23.

Kling, R. (2007), "What is social informatics and why does it matter?", *The Information Society*, Vol. 23 No. 4, pp. 205-220.

Knapp, K.J., Marshall, T.E., Rainer, R.K., Jr and Ford, F.N. (2007), "Information security effectiveness: conceptualization and validation of a theory", *International Journal of Information Security and Privacy*, Vol. 1 No. 2, pp. 37-60.

Lebek, B., Uffen, J., Neumann, M., Hohler, B. and H. Breitner, M. (2014), "Information security awareness and behavior: a theory-based literature review", *Management Research Review*, Vol. 37 No. 12, pp. 1049-1092.

Liberati, A., Altman, D.G., Tetzlaff, J., Mulrow, C., Gøtzsche, P.C., Ioannidis, J.P., Clarke, M., Devereaux, P.J., Kleijnen, J. and Moher, D. (2009), "The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration", *PLoS Medicine*, Vol. 6 No. 7, p. e1000100.

Maidabino, A.A.A. and Zainab, A.N. (2012), "A holistic approach to collection security implementation in university libraries", *Library Collections, Acquisitions, and Technical Services*, Vol. 36 Nos 3/4, pp. 107-120.

Martins, N. and Da Veiga, A. (2015), "An information security culture model validated with structural equation modelling", *In Proceedings of the Ninth International Symposium on Human Aspects of Information Security and Assurance (HAISA)*, Plymouth University, Plymouth, pp. 11-21.

Martins, A. and Eloff, J. (2002), "Information security culture", in Ghonaimy, M.A., El-Hadidi, M.T. and Aslan, K.H. (Eds), *Security in the Information Society*, Springer, Boston, pp. 203-214.

Masrek, M.N., Harun, Q.A. and Zaini, M.K. (2018a), "Assessing the information security culture in a government context: the case of a developing country", *International Journal of Civil Engineering and Technology*, Vol. 9 No. 8, pp. 96-112.

Masrek, M.N., Harun, Q.A. and Zaini, M.K. (2018b), "The development of an information security culture scale for the Malaysian public organization", *International Journal of Mechanical Engineering and Technology*, Vol. 9 No. 7, pp. 1255-1267.

Metalidou, E., Marinagi, C., Trivellas, P., Eberhagen, N., Skourlas, C. and Giannakopoulos, G. (2014), "The human factor of information security: unintentional damage perspective", *Procedia - Social and Behavioral Sciences*, Vol. 147, pp. 424-428.

Miller, M.B. (2009), "Coefficient alpha: a basic introduction from the perspectives of classical test theory and structural equation modelling", *Structural Equation Modeling: A Multidisciplinary Journal*, Vol. 2 No. 3, pp. 255-273.

Moher, D., Liberati, A., Tetzlaff, J. and Altman, D.G. (2009), "Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement", *Annals of Internal Medicine*, Vol. 151 No. 4, pp. 264-269.

Mokwetli, M. and Zuva, T. (2018), "Adoption of the ICT security culture in SMME's in the Gauteng province, South Africa", in *2018 International Conference on Advances in Big Data, Computing and Data Communication Systems (icABCD)*, Curran Associates, New York, NY, pp. 1-7.

Moody, G.D., Siponen, M. and Pahnila, S. (2018), "Toward a unified model of information security policy compliance", *MIS Quarterly*, Vol. 42 No. 1, pp. 285-311.

Nævestad, T.O., Meyer, S.F. and Honerud, J.H. (2018), "Organizational information security culture in critical infrastructure: developing and testing a scale and its relationships to other measures of information security", in Haugen, Barros, S.A., van Guijk, C., Kongsvik, T. and Vinnem, J.E. (Eds), *Safety and Reliability–Safe Societies in a Changing World*, CRC Press, London, pp. 3021-3029.

Neuman, L.W. (2014), *Social Research Methods*, Pearson Education Limited, London.

Nunnally, J.C. and Bernstein, I.H. (1994), *Psychometric Theory*, 3rd ed., McGraw-Hill, New York, NY.

Parsons, K., McCormac, A., Butavicius, M., Pattinson, M. and Jerram, C. (2014), "Determining employee awareness using the human aspects of information security questionnaire (HAIS-Q)", *Computers and Security*, Vol. 42, pp. 165-176.

Parsons, K.M., Young, E., Butavicius, M.A., McCormac, A., Pattinson, M.R. and Jerram, C. (2015), "The influence of organizational information security culture on information security decision making", *Journal of Cognitive Engineering and Decision Making*, Vol. 9 No. 2, pp. 117-129.

Paulsen, C. and Coulson, T. (2011), "Beyond awareness: using business intelligence to create a culture of information security", *Communications of the IIMA*, Vol. 11 No. 3, pp. 35-54.

Ramlall, I. (2016), "Applied structural equation modelling for researchers and practitioners: using R and stata for behavioural research", available at: https://books.google.si/books?id=YzGwDQAAQBAJ (accessed 19 June 2019).

Rančigaj, K. and Lobnikar, B. (2012), "Vedenjski vidiki zagotavljanja informacijske varnosti: pomen upravljanja informacijske varnostne culture", in Bernik, I. and Meško, G. (Eds), *Konferenca Informacijska Varnost: odgovori na Sodobne Izzive*, Fakulteta za varnostne vede, Ljubljana, pp. 1-12.

Rocha Flores, W. and Ekstedt, M. (2016), "Shaping intention to resist social engineering through transformational leadership, information security culture and awareness", *Computers and Security*, Vol. 59, pp. 26-44.

Roer, K. (2015), *Build a Security Culture*, IT Governance Ltd., Ely, Cambridgeshire.

Rubin, D.B. (2004), *Multiple Imputation for Nonresponse in Surveys*, John Wiley and Sons, Hoboken (NJ).

Saris, W.E. (1995), *The Multitratit-Multimethod Approach to Evaluate Measurement Instruments*, Eotvos University Press, Budapest.

Schlienger, T. and Teufel, S. (2003), "Information security culture - from analysis to change", *South African Computer Journal*, No. 31, pp. 46-52.

Schoenfeldt, L.F. (1984), *Psychometric Properties of Organizational Research Instruments. Method and Analysis in Organizational Research*, Reston Publishing, Reston, VA.

Singh, A.N., Gupta, M.P. and Ojha, A. (2014), "Identifying factors of 'organizational information security management", *Journal of Enterprise Information Management*, Vol. 27 No. 5, pp. 644-667.

Tabachnick, B.G. and Fidell, L.S. (2014), *Using Multivariate Statistics*, 6th ed., Pearson Education, Harlow.

Thompson, B. (2004), *Exploratory and Confirmatory Factor Analysis: Understanding Concepts and Applications*, American Psychological Association, Washington, DC.

Tinsley, H.E. and Tinsley, D.J. (1987), "Uses of factor analysis in counseling psychology research", *Journal of Counseling Psychology*, Vol. 34 No. 4, pp. 414-424.

Trochim, W.M. (2006), "Research methods knowledge base", available at: https://socialresearchmethods.net/kb/constval.php (accessed 19 June 2019).

Tsohou, A., Karyda, M. and Kokolakis, S. (2015), "Analyzing the role of cognitive and cultural biases in the internalization of information security policies: recommendations for information security awareness programs", *Computers and Security*, Vol. 52, pp. 128-141.

Yong, A.G. and Pearce, S. (2013), "A beginner's guide to factor analysis: focusing on exploratory factor analysis", *Tutorials in Quantitative Methods for Psychology*, Vol. 9 No. 2, pp. 79-94.

About the authors

Špela Orehek is a young researcher at the Center for Methodology and Informatics (University of Ljubljana, Faculty of Social Sciences) and a PhD student of Statistics. She is currently involved into issues of conceptualization and measurement of information security culture and human factors in security in general.

Gregor Petrič, PhD is a chair of the Center for Methodology and Informatics and a Full Professor of Social Informatics at the Faculty of Social Sciences (University of Ljubljana). His research interests lies in methodological issues in investigating ICT-related social phenomena, security culture, management of online communities, e-health and online interactions. Gregor Petrič is the corresponding author and can be contacted at: gregor.petric@fdv.uni-lj.si