Proposal No. 830929
Call H2020-SU-ICT-03-2018

Project start: 1 February 2019
Project duration: 42 months

Cyber Security for Europe

D9.13

Awareness effectiveness study

Document Identification	
Due date	31 January 2021
Submission date	29 January 2021
Revision	1.0

Related WP	WP9	Dissemination Level	PU
Lead Participant	NTNU	Lead Author	Sunil Chaudhary (NTNU)
Contributing Beneficiaries	NTNU	Related Deliverables	D9.18, D9.26

Abstract: This report proposes metrics for the evaluation of a cybersecurity awareness programme. In order to do so, it utilises a systematic literature review. It reviews 27 papers (selected after multiple rounds of screening) that have evaluated cybersecurity awareness mainly to extract two types of data from them i.e., what factors did the paper measure, and how did it measure those factors. After the analysis of the gathered data, it proposes four types of indicators (i.e., impact, sustainability, accessibility, and monitoring) and their respective measuring methods for the complete evaluation of a cybersecurity awareness programme. More importantly, while suggesting measuring methods, it takes into account the criteria for good metrics.

This document is issued within the CyberSec4Europe project. This project has received funding from the European Union's Horizon 2020 Programme under grant agreement no. 830929.

Executive Summary

Evaluation is an important activity of the review phase of a cybersecurity awareness programme. It provides invaluable information for the continuous improvement of the programme. More importantly, it offers an insight into the effectiveness of the programme on the audience and organisation. In addition, it provides the information required by top management or the sponsor to make a decision on whether to invest in the programme or not. Therefore, in this report, we have proposed evaluation metrics that can help to get as inclusive, complete, and unbiased results as possible.

In order to propose the evaluation metrics, we used a systematic literature review. Initially, we selected relevant papers that have analysed or proposed methods to evaluate a cybersecurity awareness program. This is followed by their review mainly to extract two types of data, which are: what factors did they measure and how did they measure those factors. We analysed the gathered data from different perspectives. Our findings revealed that these papers measured the following factors: behaviour, attitude, knowledge, interest, reachability, touchability, value-added, usability, and overall feedback. Further, in order to do so they mainly used the following methods: survey, test, passive data, face-to face-interaction, and observation. Based on these findings in along with criteria for good metrics and the European Literacy Policy Network's four indicators (i.e., impact, sustainability, accessibility, and monitoring) for awareness evaluation, we have designed and proposed evaluation metrics for cybersecurity awareness. Our proposition provides factors to be measured and their respective measurement methods in order to realise each of the aforementioned four indicators.

We believe that our proposition (metrics) will contribute to making the evaluation process of cybersecurity awareness more systematic, replicable and more importantly help to produce more accurate and trustworthy results.

Document information

Contributors

Name	Partner
Sunil Chaudhary	NTNU
Vasileios Gkioulos	NTNU

Reviewers

Name	Partner
Pasquale Annicchino	ARCH
Jozef Vyskoc	VAF
David Goodman	TDL

History

Version	Date	Authors	Comment
0.01	2020-11-15	Sunil Chaudhary	1st Draft
0.02	2021-01-26	Sunil Chaudhary	2nd Draft
0.03	2021-01-29	Sunil Chaudhary	3rd Draft
1.0	2021-01-29	Ahad Niknia	Final check and preparation for submission

Table of Contents

List of Figures

List of Tables

List of Acronyms

E	**ENISA**	European Union Agency for Cybersecurity
K	**KPI**	Key Performance Indicator
Q	**QR**	Quick Response
U	**URL**	Uniform Resource Locator

1 Introduction

Evaluating and reviewing the effectiveness of cybersecurity awareness programmes is an integral and challenging task, as such activities are most effective when performed iteratively, and focused on continuous improvement. Thus, the review phase of a cybersecurity awareness programme aims to evaluate the effectiveness of the undertaken iteration according to a set of pre-defined metrics, demonstrating in this way the achieved return on investment. Additionally, this phase also facilitates the assessment of the suitability and the necessary enhancements to future iterations (e.g., weaknesses in content quality, delivery channels, and others) to make it more relevant and effective for the target audience, and to optimise it for the organisation itself.

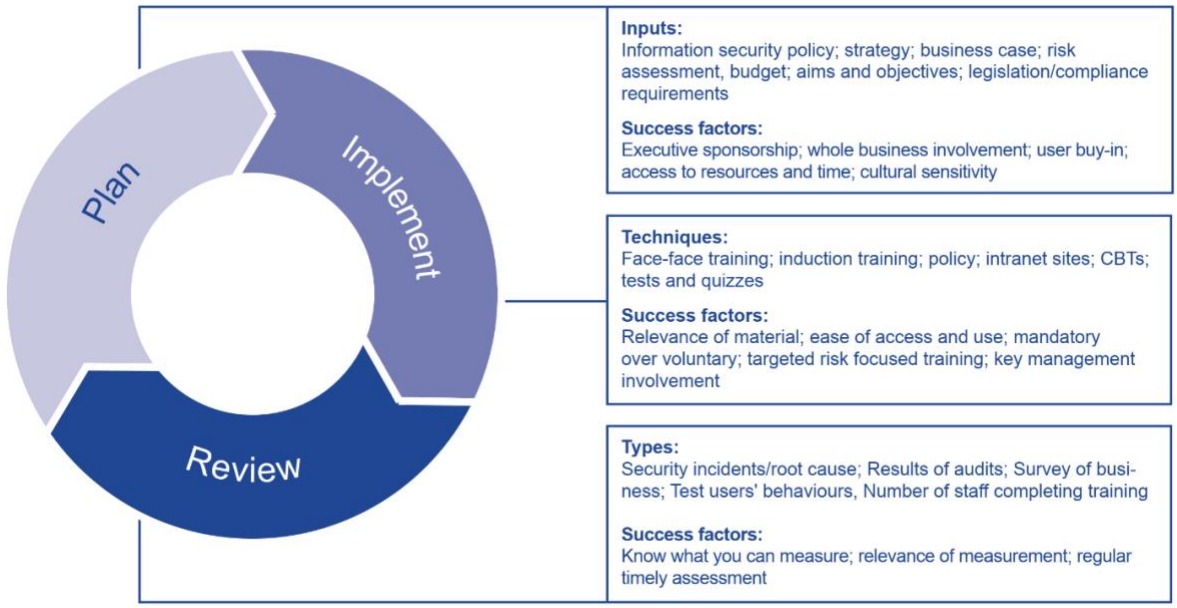

Figure 1: Iterative approach to security awareness [1]

A way could be to evaluate whether there is an expected rise in the cybersecurity awareness level of the audience or not. Even within this, there are various classifications of the cybersecurity awareness level. Kruse & Pankey [2] broadly categorised the cybersecurity awareness of an audience into five levels (shown in Figure 2):

i. *Blissfully unaware:* possesses little recognition or acceptance for most cyber threats
ii. *Consciously incompetent:* avoids risky behaviour, even if it causes productivity loss
iii. *Compliant:* is aware of the security risks identified (e.g., in the organisational policy) and takes action as mentioned in the policy
iv. *Risk aware:* considers security risks while performing their (e.g., organisational) duties, but uncertain of all mitigation actions, and
v. *Competent & practised:* recognises and mitigates security risks when performing their (e.g., organisational) duties

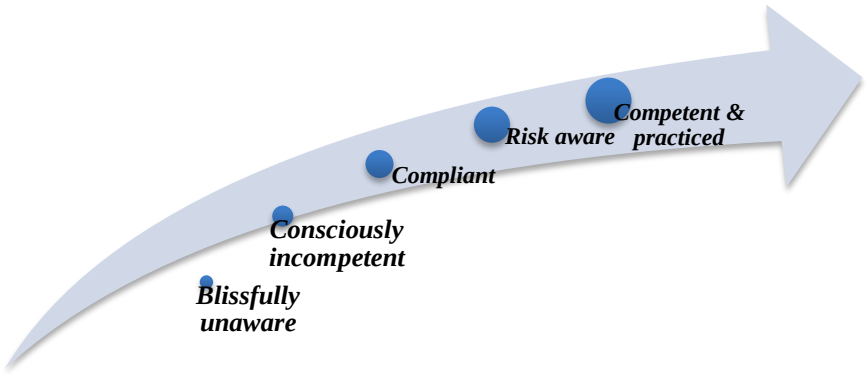

Figure 2: Cybersecurity awareness level

Similarly, the Hewlett Packard Enterprise Awareness Maturity Curve (shown in Figure 3) [3] classified the cybersecurity awareness level of an audience into:

i. *Unconscious incompetence:* does not understand cybersecurity risks and tasks and does not necessarily recognise the deficit
ii. *Conscious incompetence:* does not understand cybersecurity risks and tasks, however, is aware of the knowledge and skill gap in them as well as accepting the importance of acquiring this knowledge and skill
iii. *Conscious competence:* possesses knowledge and skill of cybersecurity risks and tasks; however, performing the tasks require practice, conscious thought, and hard work), and
iv. *Unconscious competence:* is adequately prepared for security risks and tasks and can perform the security tasks while executing another task

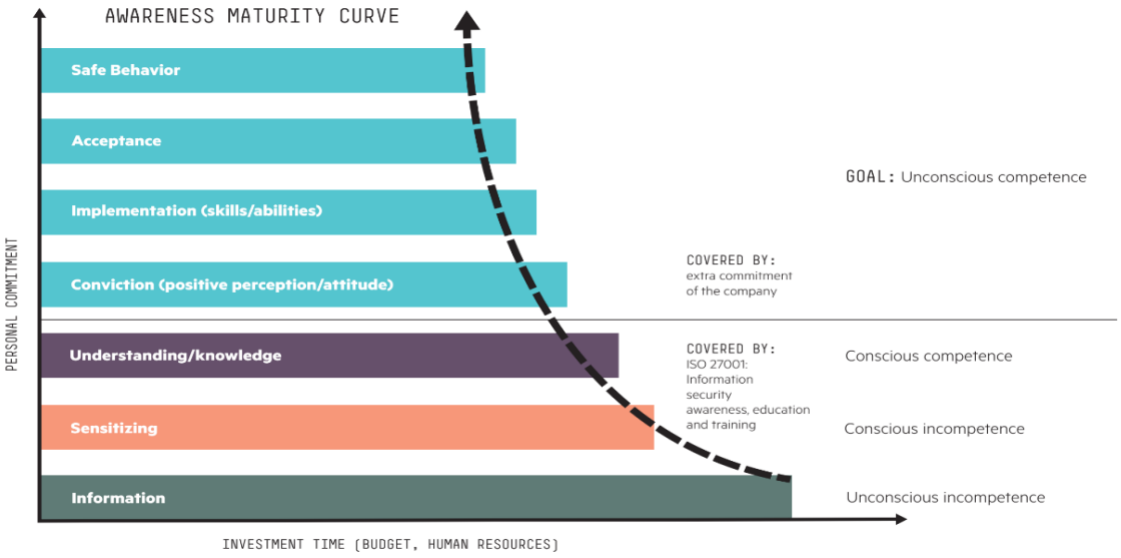

Figure 3: Hewlett Packard Enterprise Awareness Maturity Curve [4]

In both of the aforementioned classifications, ideally, the ultimate goal is to achieve the highest awareness level, i.e., raise the awareness to *competent & practised*, or *unconscious competence level*. To put it simply, the recommended security behaviours should occur automatically in an individual when performing personal and professional activities. But this goal may also vary with the needs and situations of the audience group or organisation, for example, an organisation may expect its employees to be able to understand the risks and compliance to the organisational policies (i.e., compliance level) from a cybersecurity awareness programme.

Over the last few years a number of measurements, broadly categorised as *output* and *outcome* of the programme [5], have emerged that can be used as indicators for the effectiveness of cybersecurity awareness programmes. Many past studies depend on measuring and assessing the followings to evaluate the effectiveness of cybersecurity awareness programme [6]:

i) the audience interest towards cybersecurity awareness programme generally quantified in terms of the number of participants,
ii) the reduction in the cybersecurity incidents occurred after the programme, and
iii) the change in the audience's perception, knowledge, attitude, and behaviour

Although the first parameter is simple and demonstrates the audience's satisfaction or dissatisfaction from a cybersecurity awareness programme, it does not convey whether the awareness programme made any real difference in practice. Similarly, the second parameter cannot confirm whether the improvement in cybersecurity incidents is as a result of a cybersecurity awareness programme or has occurred simply due to a decrease in cyberattacks impacting the audience after the implementation of a better firewall and network protection. The third parameter is complex but is the most relevant. It measures and assesses the changes in the security knowledge, attitude and behaviour of the audience.

Evaluation of security knowledge, attitude and behaviour uses both subjective (e.g., ask the audience about their experience), and objective (e.g., ask the audiences to do something) methods [7]. Measuring the knowledge and comprehension of security is conducted through, for example, online quizzes that can reveal whether or not people know and understand the risk [8]. Similarly, measuring attitude is conducted through, for example, anonymous surveys on why people take risky actions. But measuring the behavioural change is not simple and is performed mostly using indirect measurement, for example, self-reporting and survey. A few studies make simulated attacks on an audience and system data (i.e., log data) in order to understand the audience's security knowledge, attitude, and behaviour from their responses. Other ways include investigating triggers and motivations– two key components widely acknowledged as necessary for behaviour change to occur, that are also used to realise behaviour change [8].

Although the aforementioned parameters and their evaluation techniques are relevant, there exists no commonly agreed and understood standards on what constitutes an effective and successful cybersecurity awareness programme [9], thus hindering the evaluation process. This may have happened due to a logic that different audiences have varying needs and situations for a cybersecurity awareness programme, thus their intention for evaluation cannot be captured by metrics valid for all [10]. Ironically, this lack of metrics has become a major reason for organisations in either to not making any provision to measure the effectiveness of the programme or to evaluate merely based on the programme's outreach [10]. Without a proper evaluation, a mature cybersecurity awareness programme is presumably unachievable and above all its failure is inevitable. This brings forth a need for standard measurements for the effectiveness of cybersecurity awareness programmes. Therefore, the main objective of this report is to define the right metrics for the evaluation of a cybersecurity awareness programme and corresponding methods that provide information on how well the metrics have been met. Defining such metrics will help to reduce the ambiguity in the evaluation of a cybersecurity awareness programme by indicating the priorities that need to be focused on. Moreover, they will help to assess the aspects of the programme and identify what has been successful and what has not as well as what has been required improvement or updated. In order to define the metrics, we used a systematic literature review.

Through this proposition (evaluation metrics), we intend to make the evaluation process of the cybersecurity awareness programme as inclusive, complete, and unbiased as possible and more importantly make it *replicable* so that everyone should be able to conduct the same evaluation and get similar results. We believe that this will help cybersecurity awareness professionals assess their implementation to get more accurate and trustworthy findings for the future update and adjustment of their awareness programme.

2 Related works

Indeed, there are some major works that proposed metrics and techniques for the evaluation of a cybersecurity awareness programme. A survey report by the European Union Agency for Cybersecurity (ENISA) [1] [11] found that in general there are four main approaches, each with different performance indicators, used by organisations for the assessment of the effectiveness of cybersecurity awareness activities. Details of these four approaches are presented in Table 1. Most organisations utilise a blend of these approaches for the purpose of assessment and make their decisions based on the overall picture rather than on a single measure. Along with that, the latter work [11] also mentions that as the needs and situation of target groups differ greatly so should their evaluation metrics. Thus, it provides 71 key performance indicators (KPIs) and suggests considering different layers (i.e., business layer, service layer, and operational layer) and dimensions (i.e., planning, managing, and evaluating) while identifying the evaluation metrics and KPIs for the evaluation of the cybersecurity awareness of an organisation. Further, it recommends making use of industry-standard performance management models, such as *Balanced Scorecard* or *Six Sigma* to define performance targets and measurements.

Approach	Description	Performance indicator
Process improvement	Measures the effort invested to conduct the programme (e.g., development, dissemination, and deployment) and has no linked to the end result, i.e., whether security has improved or not. **Advantage** • Easy to define and to gather. **Disadvantage** • Provides only indirect comfort.	• Counts the main security risks or technology platforms covered • Counts staff reached • Cost of delivery (time and expenses invested per person) • Relevancy of the awareness material (the frequency with which it is updated) • Staff feedback on the awareness impact (use survey)
Attack resistance	Measures how resistant the staff is to a potential attack. **Advantage** • Provides direct evidence of the actual state of staff awareness. • Important to impress top management/sponsor and receive support and commitment to the programme. **Disadvantage** • Many attack scenarios and all of them cannot be tested. • Simulated test can be relatively expensive to set up.	• Staff ability to recognise attacks (using survey, quiz, or computer-based test) • Staff susceptibility to falling prey to attacks (using simulated attacks)
Efficiency and effectiveness	Measures the actual experience of security incidents within the organisation. **Advantage** • Easy and inexpensive to collect data. • Statistics are usually of interest to senior management. **Disadvantage** • Does not provide a true reflection of security awareness (low-security incidents can happen due to other reasons)	• Extent of: ○ security incidents (number and cost of security incidents), ○ downtime (availability of systems is critical), and ○ most severe incidents (a proportion of the total number of serious incidents) caused due to human behaviour
Internal protections	Measures secure behaviour results due to awareness. **Advantage** • Provide direct evidence of staff security behaviours.	• Extent to which security is incorporated into the development and acquisition of systems (measured by the review of business cases and requirements specifications).

	Disadvantage
	• Measure is quite specific to the behaviour it is measuring.
		• Extent to which data files are protected
		  ○ measured by the review of malware infection as shown by anti-virus activities or statistics
		  ○ measured by the report on visits to inappropriate materials or unauthorised software (from scanning tools)

Table 1: Cybersecurity awareness evaluation metrics [1]

Last but not least, Manifavas et al. [12] recommend 12 quantitative metrics for the evaluation of the cybersecurity awareness of an organisation, shown in Table 2. Furthermore, it demonstrates a method to assign a weight to metrics. In addition, it proposes the cost of implementing and running the cybersecurity awareness programme (cost-benefit analysis) as a part of the evaluation process. The effectiveness of an awareness programme is determined by the weighted summation of the value of its underlying metrics and the summation of their (i.e., metrics) costs.

Approach	Description	Performance indicator
Surveys	• Questionnaire-based survey on technical and security policy issues	• Statistical analysis of monthly survey (conducted in the different division of the organisation) and annual survey (conducted in the whole organisation)
Awareness /security day	• Direct communication with employees to get their feedback.	• Statistical analysis of security day attendance
Independent observation	• Silent observation of employees' security behaviours	• Statistical analysis of unsuccessful mock phishing attacks, and new threats bulletins' readership
Audit department reports	• Security awareness related incidents identified by audits should decline	• Count of security incidents caused due to employees' behaviour identified by the audit department
Risk department reports	• Risk identified during the previous assessment should reduce throughout time	• Count of security issues occurred due to employees' behaviour identified by the risk department

Security incidents	• Volume of security incidents occurred	• Number of employees who have caused at least one security incident
		  ○ due to their non-secure behaviour (out of the total number of employees)
		  ○ that falls within their responsibility but occurred due to their failure to identify the threat (out of the total number of employees)

Awareness sessions (workshops)	• Post session feedback from employee	• Statistical analysis of session attendance and effectiveness
Information security website	• Employees' interest in the awareness programme	• Statistical analysis of information security website visit

e-Learning	• Reachability of the awareness programme and the employees' interest in it	• Statistical analysis of e-learning programme visits, registrations, and completion
Emails	• Employee's interest in the awareness programme (link can be provided for follow up information)	• Statistical analysis of email views
iNotices	• Employees' interest in the awareness programme (link can be provided for follow up information)	• Statistical analysis of iNotice reading
Posters	• Independent observations, combined with electronic means, e.g., Quick Response (QR) code to additional resources, or Uniform Resource Locator (URL) from where the poster can be downloaded	• Statistical analysis of poster downloads

Table 2: Evaluation metrics and measuring parameters [12]

All of the aforementioned work though does not dismiss the value of qualitative methods in the evaluation process, as they all emphasise the use of quantitative methods. This is mainly because benefits that are identified but cannot be measured with quantitative values may mean less to senior management [11]. Moreover, all of them primarily focus on the evaluation of cybersecurity awareness in the organisation.

Among the three papers mentioned, the two from ENISA [1] [11] propose proper metrics for the evaluation purpose. They include both formative and summative evaluations [13] as parts of the overall evaluation of cybersecurity awareness. The work by Manifavas et al. [12] suggests what to measure and how to measure but in incoherent ways. Although these works are useful, we believe that they disregard certain important aspects that we have included and will discuss in section 5.3.

3 Research methodology

We have followed the structure of a literature review suggested by Webster & Watson [13], as shown in Figure 4.

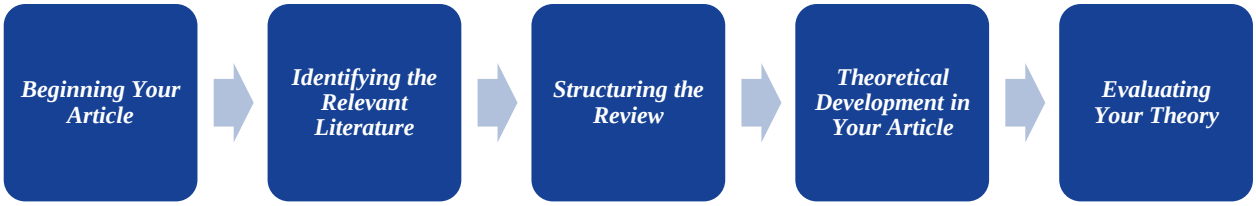

Figure 4: Structure of a literature review [13]

Initially, we have stated the problem that will be addressed by this research work and also established its relevance. This is followed by a description of some related works and their limitations, and more importantly how this research work contributes to mitigating those limitations.

In order to identify relevant literature on the topic, we used the search service on Google Scholar and Microsoft Academic. Both are freely accessible web search engines indexing an array of scholarly materials including most peer-reviewed academic journals, books, and conference papers, theses and dissertations, technical reports, and other scholarly literature published in different digital libraries and databases. No doubt, it would be exhausting to perform search and screening operations on various digital libraries, university academic repositories, and others (e.g., ResearchGate, Academia). More importantly, it is a very challenging task to decide which digital libraries and databases to include for the study and know if they will result in relevant literature or not. But it became relatively easier and more convenient to collect a large number of relevant literature studies merely by searching these two search engines, as it would otherwise have required performing search operations on different databases independently.

We used "security+ awareness+ effectiveness" as the search keyword string, where '+' is an "AND' operator. Before selecting this keyword string, we performed a trial with other keywords like "cybersecurity + awareness+ effectiveness", "cyber-security + awareness+ effectiveness", "cyber security + awareness+ effectiveness", "information security + awareness + effectiveness", and "Internet security + awareness+ effectiveness". But by using these keywords, we did not find relevant literature showing up in the top result pages.

After screening 150 results in Google Scholar Citation (until result page 15, although relevant papers stopped appearing after result page 9) and 200 results in Microsoft Academic (result page 20 although relevant papers stopped appearing after result page 18) based on their abstract and keywords, we downloaded 65 papers. In the case of search results common to both of the search engines, the download was just made from one of them. The downloaded papers were thoroughly read in the second round of screening to determine how relevant the papers were to the research topic and more specifically to our research objective, after which we finalised 27 papers for the literature review. Papers that did not include answers to the two main questions, i.e., what to measure, and how to measure to evaluate a cybersecurity awareness programme, were excluded during the screening process. We also excluded books, theses, reports, and papers in languages other than English since the working language for this project is English. However, we have not defined exclusion criteria for the year of publication, publisher, and author affiliation. Raising awareness, in general, has existed for a long time, and now in the technology era, things have changed but not significantly. Many old and traditional methods used for raising awareness are still relevant with a little modification; for example, we still use posters and leaflets to raise cybersecurity awareness. Therefore, we believe the year of publication cannot be considered as a criterion for the literature selection.

In order to structure the review, we were highly reliant on a tabular presentation style since it is easier to present a large amount of data in an understandable form.

The theory development is based on the European Literacy Policy Network, a well-established model designed for awareness evaluation. We have tailored it to make it applicable to cybersecurity awareness. While doing so, we have considered the findings of the literature review, and criteria for good metrics.

4 Literature review and data collection

During the review of the identified and selected papers, we primarily focused on what factors are measured or suggested to be measured to determine the effectiveness of a cybersecurity awareness programme and how are those factors are measured or suggested to be measured. All the reviewed papers and the relevant data collected from them are mentioned chronologically in Table 3.

Paper	What to measure?	How to measure?
Dodge & Ferguson (2006) [14]	• Behaviour to reduce cybersecurity risks	• Simulated attack (exercise) to get useful insights into the efficacy of an awareness programme

Kruger & Kearney (2006) [15]	• Cybersecurity knowledge gain (i.e., improvement in what you know)
	• Positive attitude towards cybersecurity (i.e., improvement in what you think)
	• Behaviour to reduce cybersecurity risks (i.e., improvement in what you do)
		• Questionnaire based survey to test knowledge, and realise attitude and behaviour (an indication of attitude and behaviour only when respondents do not lie)
		• Practical system data (more reliable and are not subjective or human independent, and easily obtained from the system administrator without making use of staff working time) to measure behaviour factors, e.g.,
		  ○ data like virus infection
		  ○ requests to visit unauthorised websites
		  ○ number of information technology incidents
		• Internal audit to test compliance behaviour

Kruger et al. (2006) [16]	• Cybersecurity knowledge gain (what does a person know?) • Positive attitude towards cybersecurity (how do they feel about the topic?) • Behaviour to reduce cybersecurity risks (what do they do?)	• Survey to determine the level of knowledge. • Interview to identify stakeholder wishes, concerns, problems, and values • System data, e.g., response to phishing emails, surfing of unauthorised websites, installation of anti-virus software, use of weak passwords, etc. to realise behaviour
Eminagaoglu et al. (2009) [17]	• Interest in awareness programme • Relevancy of awareness topics covered • Compliance of safe behaviour	• Percentage of attendees with respect to the number of expected attendees to measure interest • Percentage of relevant security topics covered with respect to expected topics to be covered to measure topics coverage • Periodic audit using i) technical method: tool-based attack to crack password used, ii) non-technical method: question answer survey, and iii) face to face meetings to evaluate compliance

Talib et al. (2010) [18]	• Effectiveness of awareness programme
	  ○ Usage of knowledge gained from awareness in practice
	  ○ Preference for learning method used in awareness programme
		• Survey using a questionnaire to evaluate effectiveness, to understand the usage of knowledge in practice and the preference of a learning method

Kruger et al. (2010) [19]	• Cybersecurity knowledge gain • Behaviour to reduce cybersecurity risks	• Test using vocabulary and scenario type questions for knowledge and behaviour evaluation

Albrechtsen & Hovden (2010) [20]	• Effect on cybersecurity awareness and behaviour • Interest in awareness programme • Satisfaction from the awareness programme • Intention to change cybersecurity behaviour • Intended change in cybersecurity attitude and behaviour	• Effect on awareness and behaviour is measured using a pre- and post- quantitative survey (which is anonymous) • Others are measured using a qualitative approach: interviews, group conversations and observation
Khan et al. (2011) [21]	• Cybersecurity knowledge gain • Positive attitude to cybersecurity • Normative belief and subjective norms in cybersecurity • Intention to make a positive change in cybersecurity behaviour • Behaviour to reduce cybersecurity risks	• Survey based questionnaire to measure knowledge gained • Improvements in attitude, normative belief, and subjective norms are measured using: o Interest in the awareness outcomes, i.e., counts of information security intranet page access, or visits to the webpage where awareness information is uploaded o Security-related helpdesk calls, i.e., counts of calls to helpdesk that contradict the purpose of awareness raising • Improvements in intention and behaviour are measured using: o Access to unauthorised online services o Simulated attacks and response, i.e., counts of responses to the simulated phishing emails sent o Security incident database, i.e., counts of security incidents reported
Wolf et al. (2011) [22]	• Behaviour to reduce cybersecurity risks	• Measuring the strength of passwords created before and after the awareness programme (focusing how to create a strong password)
Ahlan & Lubis (2011) [23]	• Improvement in user level of awareness o Relevancy and usefulness of awareness topics o Interest in participating in awareness programme o Feedback on awareness programme o Improvement in adaptability (i.e., opinion on how to determine and react efficiently to unsecure situation unexpectedly occurring) o Improvement in learnability (i.e., gain in knowledge or lessons learn from past actions improving the current security actions) o Improvement in performance (i.e., behaviour improving the effectiveness of security actions)	• Used quantitative survey to measure all
Rantos et al. (2012) [24]	• Awareness information reach target audience	• Information reach is measured by: o Percentage of people who attended an awareness session

	• Awareness information touch target audience	○ Counts of people that received a leaflet. ○ e-learning: number of attendees visiting the e-learning programme. ○ Email**:** Count of email recipients ○ iNotices: Count of people logged in ○ Website: Visit to the website (but can have repetitive visits from a small group of people) • Delivered information is absorbed is measured by: ○ e-learning**:** number of attendees registering, and completing the e-learning programme ○ Email**:** Hit counts to the link for more information in the email ○ iNotices: Hit counts to the link for more information ○ Poster: Poster downloaded from the link ○ Simulated attacks and response observation ○ Independent observations (e.g., awareness on clean desk policy, observation performed outside working hours) ○ Comparisons of pre- and post-session tests or survey result ○ Feedback forms (anonymous) suggestions and opinions) ○ Awareness session: Audience satisfaction, attendees temporarily leaving the room, constantly chatting with colleagues, or sketching on their notes are not encouraging reactions) ○ Anonymous survey and questionnaire, and interviews and focus group discussion to get attendee feedback and opinion ○ Share information and observe visit to the information ○ Attendance when it is not mandatory
Wilson & Hash (2012) [25]	• Behaviour to reduce cybersecurity risks • Support for awareness programme • Contribution by awareness programme • Interest towards awareness programme • Feedback strategy • External audit on effectiveness	• Behaviour changes are measured using these metrics: ○ decline in security incidents or violations ○ users being exposed to awareness materials increasing ○ coverage and identified needs are shrinking ○ users with significant security responsibilities being appropriately trained are increasing • Support for awareness programme is realised from: ○ Managers support and commitment for awareness programme. ○ Sufficient fund and resources for awareness ○ Support for broad distribution (dissemination channels) and posting of security awareness items • Contribution is realised based on recognition of security contributions (e.g., awards, contests). • Interest is realised by: ○ Attendance at mandatory security forums/ briefings ○ Motivation demonstrated by those playing key roles in managing/ coordinating security programme • Feedback strategies are: ○ Feedback forms/ questionnaires ○ Focus group (perspectives) ○ Selective interviews (feedback) • External audit is performed using: ○ Independent observation/ analysis (outside contractor or other third party as a part of an

		agency-initiated audit, unbiased opinion regarding effectiveness) ○ Security programme benchmarking (external view, benchmarking is normally done by experts)
Tsohou et al. (2012) [26]	• Positive attitude to cybersecurity • Behaviour to reduce cybersecurity risks • Deterrent effectiveness of security awareness programmes • Cost-benefit analysis of security awareness programmes	• Semi-structured questionnaire to realise the percentage of the audience who found the awareness event satisfactory (i.e., suitability and the importance of the issues discussed, programme organisation, and programme duration) • Participation in the awareness event • Percentage of awareness processes that were incorporated into the organisation's processes
Bauer et al. (2013) [27]	• Relevance or suitability of topics • Diffusion level of delivery methods • End user's cybersecurity behaviour change	• Survey is used to realise the suitability of topics • Log files are used to know diffusion level • Simulated attacks (phishing and USB social engineering) are used to evaluate behaviour change
Gundu & Flowerday (2013) [28]	• Assessment to know the needs • Awareness assessment ○ Cybersecurity knowledge gain ○ Positive attitude towards cybersecurity ○ Behaviour to reduce cybersecurity risks	• Online surveys, participant observation, informal interviews and document surveys for gathering data for needs assessment • Tests for knowledge, attitude and behaviour are used for awareness assessment
Labuschagne & Eloff (2014) [29]	• Audience interest in awareness programme • Cybersecurity knowledge gain	• Pre- and post- test (using questionnaires) to assess both interest and knowledge gained.
Velki et al. (2014) [30]	• Level of audience's risky behaviour • Level of audience's cybersecurity awareness • Level of audience's beliefs about cybersecurity	• All are evaluated using survey questionnaire
Manifavas et al. (2014) [12]	• Awareness information reach (i.e., information reached target audience) • Awareness information touch (i.e., target audience absorbed delivered information) • Cost of implementing and running awareness programme	• Survey is used to measure information reach • Information touch is measured using: ○ Awareness or security days communication (face-to-face feedback) ○ Independent observations (behaviour) ○ Audit department reports (count of security issues related to employees) ○ Risk department reports ○ Security incidents • Financial calculation is used to measure cost
Prah et al. (2016) [31]	• Cybersecurity knowledge gain • Positive attitude to cybersecurity • Behaviour to reduce cybersecurity risks	• Knowledge is measured using test • Attitude and behaviour changes are evaluated using open-ended questions.
Scholl et al. (2017) [32]	• Cybersecurity knowledge and competence gain • Willingness to behave securely • Behaviour to reduce cybersecurity risks • Self-responsibility to behave securely	• Standardised survey questionnaire is used to measure knowledge and competence • Qualitative interview is used to know willingness to behave securely • Appearance of security incidents, e.g., helpdesk reports or the results of virus scans are used to monitor self-responsibility • Simulated attack and its response observation (secretly) are used to realise cybersecurity behaviour
Carella et al. (2017) [33]	• Positive attitude to cybersecurity • Willingness to learn about cybersecurity	• Survey is used to evaluate attitude and willingness

	• Cybersecurity behaviour change	• Click through rate (clicked on the malicious link in email with respect to its view) is used to know behaviour change
Wahyudiwan et al. (2017) [34]	• Cybersecurity knowledge gain • Positive attitude towards cybersecurity • Behaviour to reduce cybersecurity risks	• Questionnaire based survey is used for all
Al Shamsi (2019) [35]	• Raise in cybersecurity awareness • Behaviour to reduce cybersecurity risks	• Semi-structured interview is used for all

Gundu et al. (2019) [36]	• Delivery assessment
	• Awareness assessment
	  ○ Cybersecurity knowledge gain
	  ○ Attitude to cybersecurity
	  ○ Intention to adopt cybersecurity behaviour
		• Pass and fail rates, frequency of training, and count of attendee are used for delivery assessment
		• Web-based questionnaire /survey tests, general observation, antivirus and firewall statistics, and incident logbook are used for awareness assessment

Ikhalia et al. (2019) [37]	• Usability of awareness programme
	• Behaviour to reduce cybersecurity risks
		• Usability of awareness programme is measured using:
		  ○ Usefulness of the awareness programme and user satisfaction using closed questionnaire
		  ○ Semi-structured interview to get opinion and feedback on the delivery channel
		• Laboratory experiment (paired sample t-test) is used to understand security behaviour

Tschakert & Ngamsuriyaroj (2019) [38]	• Performance (i.e., improvement in cybersecurity knowledge, attitude, and behaviour) • Confidence (i.e., learned things are useful in real life) • Satisfaction (i.e., learned things and enjoyed learning) • Preference (i.e., liked the content and delivery method used)	• Pre- and post- awareness simulated attacks are used to measure performance • Pre- and post- awareness questionnaire-based test (phishing screenshots to identify) are used to measure performance • Post-awareness questionnaire is used to realise confidence, satisfaction, and preferences

Haney & Lutters (2020) [39]	• Awareness impact measurement
	  ○ Information reaches the target audience
	  ○ Audience feedback
	  ○ Behaviour change
		• Cybersecurity information reach to the right audience is measured using:
		  ○ Attendance can be an indicator of reach. Who is attending will give an insight into those buying into the importance of the awareness programme, whether the programme is reaching the right people?
		• Audience feedback is obtained using
		  ○ Informal break room conversation
		  ○ Anonymous post-event surveys
		• Effectiveness is measured using:
		  ○ Analysis of user-generated security incidents data aggregated from multiple sources, e.g., after security awareness training regarding the sending of sensitive information via email, are the number of personal data disclosures going down? Do not just focus on where employees fall short; look at indicators of positive behaviours as well as increased reporting of suspicious emails or other security incidents to the help desk.

Table 3: List of reviewed papers and data retrieved

5 Data analysis and resulting metrics

5.1 Factors evaluated

In order to evaluate the effectiveness of a cybersecurity awareness programme, the reviewed papers measured the factors given in Table 4, and the count of papers measuring them are shown in Figure 5.

Measured factor	Paper
Improvement in cybersecurity **behaviour** resulted from participating in an awareness programme. This has been expressed as follows: • Reduction in cybersecurity risky behaviour • Promotion of the best practices and compliance of safe behaviour • Positive effect on cybersecurity behaviour • Intended change in cybersecurity behaviour • Intention-in-action to change cybersecurity behaviour • Improvement in performance (i.e., change in behaviour improving the effectiveness of security actions) • Deterrent effectiveness (i.e., discourage risky actions) • Level of audience's risky behaviour • Self-responsibility to behave securely	References [14], [15], [16], [19], [20], [21], [22], [23], [25], [26], [27], [28], [30], [31], [32], [33], [34], [35], [36], [37], [38], [39]
Positive changes in the cybersecurity **attitude** of the audience resulted from participating in a cybersecurity awareness programme. This has been expressed as follows: • Developed positive attitude towards cybersecurity • Intended change in cybersecurity attitude • Normative belief and subjective norms towards cybersecurity • Improvement in adaptability (i.e., opinion on how to determine and react efficiently to unsecure situation occurring unexpectedly) • Level of audience beliefs in cybersecurity • Willingness to behave securely • Willingness to learn about cybersecurity • Intention in words to make positive changes in cybersecurity behaviour	References [14] [15], [16], [20], [21], [23], [26], [28], [30], [31], [32], [33], [34], [35], [36], [38]
Cybersecurity **knowledge and competence** gained by participating in a cybersecurity awareness programme. This has been expressed as follows: • Cybersecurity knowledge gained • Level of audience cybersecurity awareness • Cybersecurity knowledge and competence gained • Improvement in learnability (i.e., gain in knowledge or learn from past actions improving the current security actions)	References [14] [15], [16], [19], [21], [23], [28], [31], [29], [30], [32], [34], [35], [36], [38]
Audience, organiser, and management/sponsor **interest** in a cybersecurity awareness programme. This has been expressed as follows: • Audience interest in an awareness programme • Audience interest to participate in an awareness programme • Motivation demonstrated by the organiser of an awareness programme • Managers or sponsors support and commitment for an awareness programme	References [17], [20], [23], [25], [26], [29]
Reachability of an awareness programme, i.e., information has reached the right audience. This has been expressed as follows: • Awareness information reached the target audience • Diffusion level of delivery methods	References [24], [27], [12], [39]
Touchability of an awareness programme, i.e., information is perceived positively by the right audience. This has been expressed as follows: • Awareness information touched the target audience • The target audience absorbed the delivered information	References [24], [12]

Value added by an awareness programme, i.e., economical or other benefits. This has been expressed as follows: • Contribution by an awareness programme • Cost-benefit analysis of an awareness programme • Cost of implementing and running an awareness programme (cost saving)	References [26], [25], [12]
Usability of topics covered, learning methods used, and awareness programme organised. This has been expressed as follows: • Relevancy of awareness topics covered • Relevancy and usefulness of awareness topics • Relevancy or suitability of topics • Usage of knowledge gained from awareness in practice • Gain in confidence (i.e., learned things that are useful in real life) • Preference for learning method used in awareness programme • Preference (i.e., liked the content and delivery method used) • Satisfaction from an awareness programme • Satisfaction (i.e., learned things and enjoyed learning) • Delivery assessment • Usability of an awareness programme	References [17], [18], [20], [23], [27], [26], [36], [37], [38]
Overall feedback on an awareness programme. This has been expressed as follows: • Feedback on awareness programme • Feedback strategy • Audience feedback	References [23], [25], [28], [39]

Table 4: Factors measured by the reviewed papers

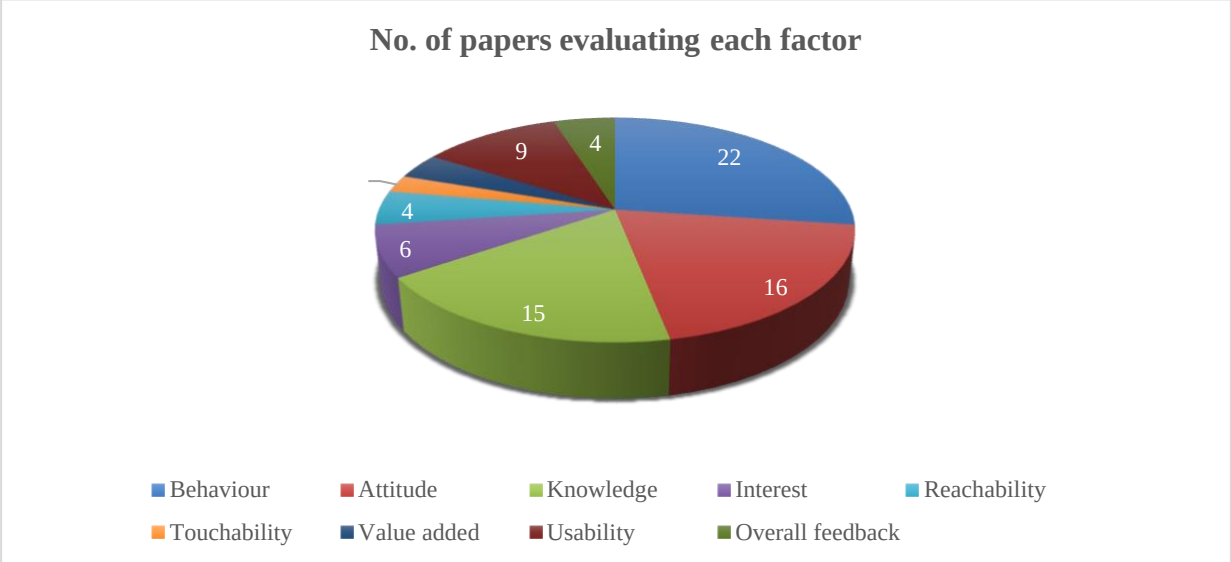

Figure 5: Factors measured by the reviewed papers

5.2 Methods used for evaluation

The methods utilised to evaluate each factor are presented in Table 5. Basically, these methods can be classified into two types [40]:

- **intrusive** – participant's normal behaviour is consciously disrupted by the evaluation processes, and
- **non-intrusive** – participant's normal behaviour is not consciously disrupted by the evaluation processed

Among the intrusive methods surveyed, interview, focus group discussion, and test have been used for the evaluation processes, and similarly among non-intrusive methods system and log data, observation, simulated and other attacks (secret) have been utilised for the evaluation processes (as shown in Figure 6).

Measured factor	Measurement method
Behaviour	**Intrusive technique** • *Simulated attack* and its response observation (pre- and post- attacks), e.g., count of responses to the simulated phishing emails sent, or click-through rate of malicious link in an email (secret attacks but both legal and ethical considerations are musts) • Questionnaire-based *survey* (qualitative; open-ended questions; pre- and post-survey), *face-to-face meeting*, semi-structured *interview*, and *group discussion* (valuable only when respondents do not lie) • *Laboratory experiment* and a silent observation of participant's actions (conscious audience so the results may not exactly depict that of the subconscious and natural behaviour) • Web-based *test* by using vocabulary and scenario type questions (conscious audience so may attempt to perform better)

	Non-intrusive technique
	• *Practical system data* to measure an increment in compliance of the best security practices, or a reduction in risky behaviours (more reliable and not subjective or human dependent, and easy as well as inexpensive to obtain)
	  ○ security incidents or violations reported, e.g., virus infection incidents (from incidents logbook)
	  ○ request to visit or access and surfing of unauthorised online services and websites
	  ○ use of weak passwords
	  ○ installation of anti-virus software
	  ○ sending of sensitive information via email
	  ○ personal data disclosure or breach
	• *Tool-based attack*, for example, to crack a password and measure strength of passwords created before and after the awareness programme (is not possible for every cyber threat)
	• *Silent observation* of compliance (in organisation, preferably, after work hours)

Attitude	**Intrusive technique** • Questionnaire based *survey* (quantitative or qualitative; open-ended questions), semi-structured *interview*, and *group discussion* to know wishes, concerns, problems, values, belief, norms, and willingness of cybersecurity (valuable only when respondents do not lie) • *System data (*interest in an awareness programme), for example, counts of information security intranet page accesses, or visits to webpage where awareness information is uploaded
	Non-intrusive technique • *Security related helpdesk calls*, i.e., count of calls to helpdesk that run counter to the purpose of awareness raising • *Silent observation* of security-related activities
Knowledge and competence	• Standardised *survey* questionnaire to measure knowledge and competence • (Pre- and post-) *tests* using vocabulary and scenario type questions, e.g., phishing screenshots to identify
Interest	**Interest by audience** • Percentage of attendees (i.e., *attendance*) with respect to the expected number of attendees (if mandatory in the organisation, most employees are forced to attend, and may not represent the real interest; voluntary participation shows the real interest) • *Survey* (quantitative or qualitative) and other qualitative approaches, e.g., interviews and group discussion • *Silent observation* of participants during the session, e.g., yawning, side talking, frequency of short break taken.

	Interest by organiser • Motivation demonstrated (*observation*) by those playing key roles in managing/ coordinating cybersecurity programme
	Interest by management • Moral support and commitment by management (*observation and interview*) for an awareness programme • *Fund and resources allocated* for an awareness programme, for example, to support for distribution (i.e., use of dissemination channels) and posting of security awareness items
Reachability	**Accessibility of awareness materials** • Percentage of people who attended an awareness session • Count of people that received a leaflet. • Number of attendees visiting the programme in e-learning • Count of email recipients • Count of people logged into iNotices • Visit to the website (but can have repetitive visits from a small group of people) • Survey to know who received the awareness information
Touchability	**Self-motivated actions** • Attendance when it is not mandatory • Number of attendees who registered and completed the programme in e-learning • Hit counts to the link for more information in the email • Hit counts to the link for more information in iNotices • Posters downloaded from the link provided • Simulated attacks and response observation • Independent observations (e.g., awareness on clean desk policy, observation performed outside working hours). • Comparisons of pre- and post-session tests or survey results • Feedback forms, survey (anonymous), interviews and focus group discussions • Audience satisfaction (e.g., attendees temporarily leaving the room, constantly chatting with colleagues, or sketching on their notes are not encouraging reactions to be discouraged) • Visits to shared information • Awareness or security day communication (face-to-face feedback) • Independent observations (behaviour) • Audit and risk department reports (count of security issues related to employees) • Security incidents reported
Value added	**Non-Financial benefit** • Contribution is realised based on recognition of security contributions, e.g., counts and reputation of awards and contests won. • Percentage of awareness processes incorporated in the organisation's processes
	Financial benefit • Financial cost calculation of organising an awareness programme
Usability	**Relevant topics covered** • Percentage of relevant security topics covered (with respect to expected topics to cover) • Survey, interview and group discussion to realise the covered topics were suitable to the audience
	Delivery assessment • Post awareness survey to know the learning method was preferred by the audience • Pass and fail rates, frequency of training, and count of attendee
	Usage of knowledge in practice

	• Survey using questionnaire to know the usage of knowledge in practice
	User confidence and satisfaction • Users exposure to awareness materials is increasing • Users with significant security responsibilities being appropriately trained is increasing • Coverage and identified needs are shrinking • Survey using closed questionnaire • Satisfaction measured using qualitative approach: interviews, group conversations and observation • Post awareness questionnaire to realise confidence, satisfaction and preferences
	Usefulness of awareness programme • *Survey* using closed questionnaire • Semi-structured *interview* to realise the percentage of audience that found the organisation of the event satisfactory (i.e., suitability and importance of the issues discussed, programme organisation, and programme duration)
Overall feedback	**Feedback strategies** • Post-event *survey* that can be qualitative or quantitative (preferably anonymous) • *Feedback forms* (preferably Anonymous) • *Focus group* discussion • Selective/informal *interview* • Informal break room *conversation*

Table 5: Factors measured and their respective measurement techniques

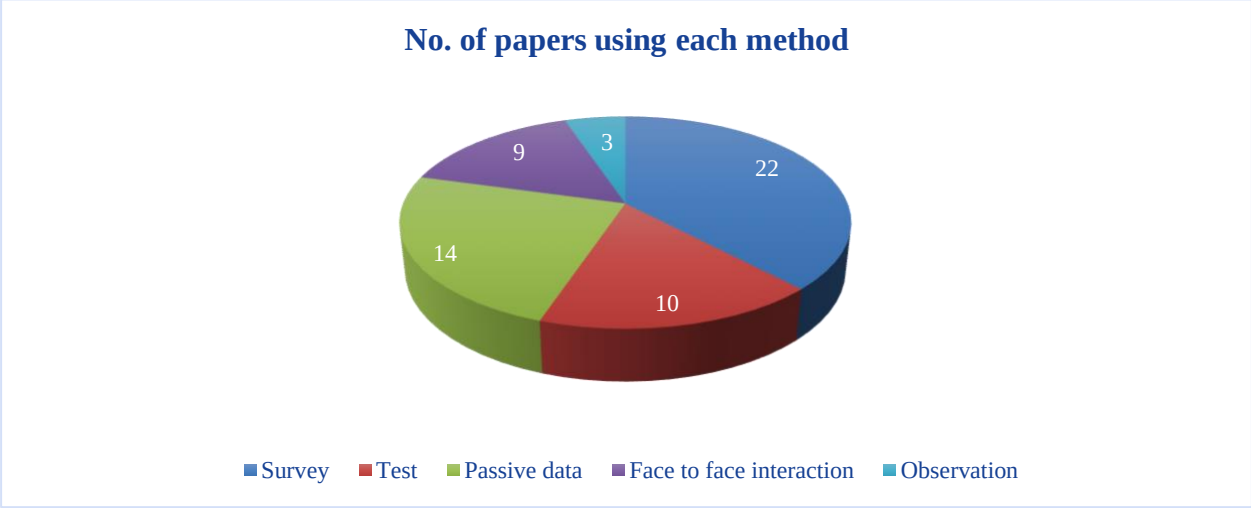

Figure 6: Methods used for the evaluation cybersecurity awareness

5.2.1 Survey

A survey using questionnaires has been found to be the most popular method for the evaluation of cybersecurity awareness programme. It has been implemented mainly to determine the impact of a cybersecurity awareness programme on cybersecurity knowledge, attitude, and behaviour of the audience. In order to do so, pre (before the awareness programme) and post (after the awareness programme) survey have been utilised. When close-ended questions are used, the results can be quantified; however, open-ended questions are presumably more suitable for the evaluation of attitude and behaviour [31]. Moreover, using a *standardised survey questionnaire* to measure knowledge and competence [32] can help to get more reliable and scientifically valid results.

In addition to these factors, post-survey has also been implemented for many other purposes, such as:

- to determine the suitability and usefulness of the covered topics
- to realise the importance of the knowledge gained in practice
- to understand the interest and willingness in participation
- to realise confidence, satisfaction, and preference of the audience
- to determine preference for learning methods
- to get overall feedback (or suggestion, opinion)

Even though it is not mandatory to have an anonymous survey, making it anonymous can encourage the participants to provide real and honest feedback. This is important since the survey result is of value as long as the participants do not lie.

5.2.2 Passive data

Passive data is collected from multiple sources, such as the audit department, the risk department, other external and internal auditors, and the helpdesk in a natural environment (i.e., participants remain unaware of the data collection for a research purpose). This data is not subjective (i.e., human independence so a separate time of the audience for the data collection is not required), and easy as well as economical to obtain. These may be some reasons why many studies utilised them to evaluate cybersecurity behaviours (both risky behaviours and best security or compliance behaviours). Some types of passive data used to measure cybersecurity behaviour are:

- Anti-virus and firewall logs
- Visit or requests to visit unauthorised services and websites
- Number of security incidents or violations reported
- Use of weak passwords
- Sending of sensitive information via email
- Count of calls to the helpdesk
- Visits or traffic to the location where awareness information is available (e.g., security intranet page, or location where awareness information is uploaded)
- Click through rate of malicious links
- Count of information security intranet page access, or visits to the webpage where awareness information has been uploaded
- Installation of security protection
- Coverage and identified needs of cybersecurity awareness are shrinking
- Frequency of awareness programme needed in the organisation
- Increase in reporting of potential cyber incidents by cyber aware people

Similarly, passive data that has been utilised to know whether the cybersecurity awareness information reached the target audience or not are:

- Count of people that received a leaflet
- Number of attendees visiting the e-learning programme
- Count of the email recipients
- Count of people logged into iNotices
- Visits to the website (but there is a risk that a small group of people may repeatedly visit the website)
- Percentage of people who attended an awareness session

The audience's interest in a cybersecurity awareness programme (or whether cybersecurity information touched the audience or not) is also determined by utilising the following passive data:

- Number of attendees registering, and completing the e-learning programme
- Hit counts to the link for more information in the email
- Hit counts to the link for more information in iNotices
- Poster downloaded from the link

- Activities like attendees temporarily leaving the room, constantly chatting with colleagues, or sketching on their notes are not encouraging reactions

As a matter of fact, utilising such data for the evaluation purpose in cybersecurity provides a more realistic outlook of the situation. This data is a part of everyday activities so participants do not need a separate notification that could make them aware and alert thus influencing their activities and data.

5.2.3 Test

Tests in two forms have been utilised for evaluation purposes, which are i) a question-based test, and ii) an attack-based test. Such tests are performed before and after a cybersecurity awareness programme and the two results are compared to know the effectiveness of a cybersecurity awareness programme. These tests are conducted mainly to evaluate cybersecurity knowledge, and in case of a simulated attack are also used to evaluate behaviour. In a question-based test (e.g., quiz or game) using standardised questions [32] comprising vocabulary and scenario type questions [19] can help to ask the right and relevant questions. Then, in an attack-based test, using a secret simulated attack, for example, sending phishing emails to the audience and observing the response can provide more realistic result. Similarly, other attack types, such as checking password strength using tools and techniques after an awareness programme on creating strong passwords can also be a test to evaluate the effectiveness of the awareness programme. However, while conducting such attacks it is mandatory to ensure that no laws and ethics are contravened.

As with the passive data approach described above, this method also provides a more realistic view of the situation. But it involves more works (like developing attacks in as natural form as possible, taking care of legal and ethical aspects, and others) and can be expensive to conduct.

5.2.4 Face-to-face interaction

Face-to-face interaction using semi-structured interviews, informal break room conversations, and focus group discussions to get audience feedback can be either targeted or generalised (e.g., suggestions, opinions, wishes, concerns, problems, and values) on a cybersecurity awareness programme. Such face-to-face interaction, for example, in a laboratory experiment can be utilised to an extent to realise audience cybersecurity knowledge, attitude and behaviour.

5.2.5 Observation

Observation has been utilised mainly to evaluate cybersecurity behaviour. This is an effective method to notice a change in audience behaviour after participating in an awareness programme; however, quantifying the outcome may not be simple.

5.3 Metrics development

But prior to metrics development, it is important to realise what constitutes a good metric. Some criteria of what constitutes good metrics, which we believe are relevant for our proposed metrics [4], are shown in Figure 7.

CRITERIA FOR GOOD METRICS	Consistently measure (i.e., no subjective criteria)
	Cheap or economical to gather (i.e., preferably automated)
	Expressed as a cardinal number or percentage
	Expressed using at least one unit of measure
	Contextually specific (i.e., relevant to decision makers so they can take action)

Figure 7: Criteria for good metrics [4]

Evaluation can be:

- *diagnostic*- a pre-assessment conducted to know an audience's existing awareness level on the topic,
- *formative*- an assessment conducted during the programme development and implementation to realise the needs and processes required to achieve the goal, and
- *summative*- a post-assessment conducted to assess the outcome of the programme and determine broader and long-term changes occurred due as a result of the programme

The diagnostic assessment followed by the summative assessment is mainly related to the outcome and impact of the programme or the declaration of the success or failure of the programme, but the formative assessment helps learn where to best put limited resources. For a complete evaluation of a cybersecurity awareness programme, all three assessments are necessary.

For the evaluation process, it is requisite to have a clear and measurable goal and objective (or expectation) from a cybersecurity awareness programme. In general, a cybersecurity awareness programme is expected to:

- communicate cybersecurity knowledge (i.e., recommended guidelines and security best practices) to the target audience,
- broaden the cybersecurity knowledge of the target audience (i.e., familiarity with guidelines and security best practices),
- bring positive changes in attitude (i.e., motivate to adopt recommended guidelines and practices) and behaviour (i.e., create a strong culture of security) in the target audience,
- gain and keep the audience and management/ sponsor trust and satisfaction, and ultimately
- minimise the number and extent of security breaches [11].

In addition, the programme must be cost-effective (or inexpensive) to conduct [11]. In other words, the success of a cybersecurity awareness programme is defined in terms of its non-financial and financial effects.

The European Literacy Policy Network [41] recommends four indicators and their measurement methods that are important for the evaluation of awareness activities. Based on this recommendation, the aforementioned criteria for good metrics, and evaluation methods utilised by the reviewed studies, we propose the metrics as shown in Table 6 for the evaluation of a cybersecurity awareness program. We believe that all these four indicators are important to be evaluated in order to know the effectiveness and success of a cybersecurity awareness program. It is possible that an organisation may not be in a situation to afford the measurement of every factor. In a situation like this, it is suggested that the organisation measure selective factors most relevant to it from each indicator rather than measuring all factors from a certain indicator while abandoning other indicators. The target audience will impact how each indicator can be measured. For example, it may be easy and economical to obtain system and log data if the target audience is the organisational staff (they are in a controlled environment) but such data may not always exist if the target audience is customers (they are in an uncontrolled environment). Moreover, while suggesting measurement/ assessment methods, we have tried to ensure that they adhere to criteria for good metrics. For example, we have emphasised a quantitative method that is non-subjective as well as quantifiable, and so making sense to the sponsor/management. Besides, we have provided multiple alternatives to measure each indicator type.

In addition, an evaluation should not be limited to what factors to measure and how to measure them but should also cover whom they have been measured against. This will help in the complete evaluation (i.e., from the perspective of all-important stakeholders like cybersecurity awareness professional, management/ sponsors, and an audience group) and at the same time provide an idea of who will participate in the evaluation process. Outcomes from the evaluation of *impact factors* and *accessibility factors* are more connected to the cybersecurity awareness professionals who are responsible for

updating and adjusting the cybersecurity awareness programme for future iterations. Whereas evaluation results of *sustainability indicators* and *monitoring indicators* are helpful for the management or sponsor in deciding whether to continue investing in the existing awareness programme or has to look for an alternative.

Indicator	Measured factor	Measurement/Assessment method
Impact indicators	Impact of awareness on: • Cybersecurity knowledge & competence • Attitude to cybersecurity • Cybersecurity behaviour	• (Pre- and post-, quantitative) question-based test to determine if the audience knows more about the issued covered by the awareness programme or not. • (Pre and post) statistical analysis of passive data to know if there is a decline in security incidents and violations, for example, o Data from audits and risk departments o Counts and severity of security incidents occurred due to staff behaviours o Other best behaviour data that can be automatically collected (e.g., anti-virus and firewall log data, and helpdesk data) • (Pre and post) simulated attack to determine if the audience understands the sense of urgency of fighting and preventing the issue or not.
Sustainability indicators	Impact of awareness in the change of: • Organisational policies • Regulatory framework • Organisational arrangement Change in top management and sponsor support and commitment for the awareness programme	• Valued-added by the awareness programme evaluation based on: o Recognition of security contributions, e.g., count and reputation of awards and contests won due to the awareness programme o Percentage of awareness processes incorporated in the organisation's policies, processes, and arrangement • Change in funding and resources allocated for the awareness programme to realise the management/ sponsor interest in the awareness programme • Cost-benefit analysis of the programme (i.e., return on investment)
Accessibility indicators	• Quality of awareness resources • Effectiveness of awareness resources	• Survey to evaluate (using closed questions/ quantitative, such as Likert scale): o relevancy of topics o delivery assessment o usage of knowledge in practice o satisfaction o usefulness of the programme • Percentage of security topics covered with respect to expected topics to be covered to know if all relevant or demanded topics are covered or not • System and log data analysis (e.g., attendance, website visit, email recipient, etc.) to determine if the target group has access to the awareness resources or not
Monitoring indicators	Interest and active participation in the programme.	Interest and active participation (touchability) evaluated using: • System and log data analysis (e.g., attendance when it is not mandatory, number of attendees who

		registered and completed the e-learning programme with respect to those who visited, hit counts to the link for more information etc.) • Post-event survey (using closed questions/ quantitative, such as Likert scale; preferably anonymous) to receive overall feedback on the awareness programme.

Table 6: Metrics for the evaluation of cybersecurity awareness

Different to the works discussed in section 2 (i.e., related works), our proposed metrics have given equal importance to the evaluation of the sustainability indicators. Sustainability can be expressed in terms of the programme outcome's ability to exist constantly by influencing organisational policies, arrangements, and regulatory framework. It can also be expressed in terms of the programme's ability to exist constantly in the organisation by becoming a part of the organisational policies and receiving abundant funds. Cybersecurity awareness is a continuous process and the evaluation process is similarly iterative: and without the evaluation of the sustainability indicator, the continuity of the process itself can become questionable. In addition, the evaluation should seek input from all of those involved and affected by the evaluation. This is possible only by ensuring that diverse viewpoints from different stakeholders are taken into account so that the results are as complete and unbiased as possible. Ironically, none of the work discussed in section 2 has clearly considered this aspect and has based their evaluation completely on the audience viewpoint.

The proposed metrics be a can guide for the evaluation process of a cybersecurity awareness programme; however, it does not answer what score is an acceptable level of awareness [12] [15]. This is an important question but is contextual and will vary depending on the target topic and audience type. For example, if the target audience is healthcare or banking staff, the only acceptable score will presumably be the maximum. Therefore, it is necessary to set a benchmark expectation on the cybersecurity awareness programme [11] [25].

6 Conclusions and future work

The evaluation of a cybersecurity awareness programme is an important activity of the post-implementation phase. Evaluation is necessary to know how effective and successful the programme was. Moreover, it provides information on which aspects of the programme requires improvement and also information used by senior management/ sponsor in deciding whether or not to invest further in the programme.

In spite of all the benefits of evaluation, there does not exist a consensus on what to measure and how to measure while evaluating a cybersecurity awareness programme. This may be because different target groups have varying needs and environments determining the content of their cybersecurity awareness programmes; so, the rationale behind an evaluation strategy cannot be captured by generalised evaluation metrics. Ironically, this lack of evaluation metrics for cybersecurity awareness has caused more harm than good; for example, many organisations and individuals either abandon the evaluation process or limit their evaluation to some weak or irrelevant factors and indicators. Therefore, in this report, we have designed and proposed evaluation metrics for cybersecurity awareness that we believe are widely applicable.

In order to do so, we performed a systematic literature review of 27 past studies that have evaluated or proposed methods to evaluate a cybersecurity awareness programme. The relevant papers were gathered after multiple rounds of screening. This is followed by a review of the gathered papers mainly to extract information on what factors past studies measured and how they measured them to evaluate or assess the effectiveness and success of a cybersecurity awareness programme. Analysis of the collected data revealed that factors measured by the past studies can be classified into behaviour, attitude, knowledge, interest, reachability, touchability, value-added, usability, and overall feedback. Among all the factors measured, behaviour, attitude and knowledge are the most popular factors. Similarly, methods used to measure these factors can be categorised into survey, test, passive data, face-to face-interaction, and observation, where survey and passive data are found to be the most popular.

Using the obtained findings, criteria for good metrics, and the European Literacy Policy Network's four indicators (i.e., impact, sustainability, accessibility, and monitoring), we have designed and proposed new metrics for the evaluation of a cybersecurity awareness programme. Our proposition provides factors to be measured and their respective measurement methods in order to realise each of the indicators.

- The *impact indicator* is realised by measuring positive changes in cybersecurity knowledge, attitude, and behaviour due to the cybersecurity awareness programme using methods like test (both question-based and simulated tests) and passive data.
- Similarly, the *sustainability indicator* is realised by measuring the changes in organisational policies, regulatory framework, and organisational arrangement due to cybersecurity awareness. Moreover, it is also realised by measuring the change in senior management and sponsor support and commitment for the cybersecurity awareness. The sustainability indicator is measured using methods like percentage of awareness outcome integrated into the organisational process, policy, and arrangements; cost-benefit analysis; and changes in funds and resources allocated for the programme.
- Next, the *accessibility indicator* is realised by measuring the quality of awareness materials and effectiveness if delivery channels using methods like survey, the percentage of relevant topics covered, and the audience's interest in the awareness programme.
- Finally, the *monitoring indicator* is realised by measuring audience interest and active participation in the awareness programme using passive data analysis and post programme survey.

We believe our proposition is inclusive of all directly affected stakeholders, i.e., management, cybersecurity awareness professionals, and target audiences. More importantly, the proposed metrics have taken into account various important aspects, such as criteria for good metrics, different stakeholder needs, and the sustainability of the programme in order to make the evaluation process inclusive, complete, and unbiased as far as possible.

Last but not least, this is the first in a series of three reports on Awareness Effectiveness. In future work, we plan to explore how these metrics can be implemented to evaluate the impact of cybersecurity awareness programme in small and medium-sized enterprises (SMEs) and supply chain. Moreover, there are other aspects like when the evaluation should be performed (i.e., timing and frequency), and how weighting can be incorporated into the proposed metrics to get an overall result of the evaluation that required to be explored.

7 References

[1] ENISA, “Information security awareness initiatives: Current practice and the measurement of success,” *ENISA*, July 2007, https://ifap.ru/library/book206.pdf.

[2] S. Kruse & B. Pankey, “Assessing the effectiveness of security awareness training,” http://www.securitymetrics.org/attachments/Metricon-6.5-Kruse.pdf.

[3] M. Beyer et al., “Awareness is only the first step: A framework for progressive engagement of staff in cyber security,” *Hewlett Packard Enterprise* , UK, 2015, https://www.riscs.org.uk/wp-content/uploads/2015/12/Awareness-is-Only-the-First-Step.pdf.

[4] A. Jaquith, "Security metrics: Replacing fear, uncertainty, and doubt", Boston, Massachusetts, United States: Addison-Wesley Professional, 2015.

[5] N. Rohlich et al., “Exploring the effectiveness of transit security awareness campaigns in the San Francisco Bay area,” *Mineta Transportation Institute*, San José, CA, USA, June 2010, https://transweb.sjsu.edu/research/Exploring-Effectiveness-Transit-Security-Awareness-Campaigns-San-Francisco-Bay-Area

[6] L. Spitzner, “Security awareness metrics,” *SANS*, https://www.sans.org/security-awareness-training/blog/security-awareness-metrics.

[7] B. Timmermans & A. Cleeremans, “How can we measure awareness? An overview of current methods,” in *M. Overgaard (Ed.), Behavioural Methods in Consciousness Research*, Oxford, UK, Oxford University Press, 2015, p. 21–46.

[8] Cybsafe, “Measuring the effectiveness of security awareness training,” *Cybsafe*, 2 November 2018, https://www.cybsafe.com/blog/measuring-the-effectiveness-of-security-awareness-training/

[9] R. Richardson, “CSI computer crime & security survey,” *CSI*, 2008, http://www.sis.pitt.edu/jjoshi/courses/IS2150/Fall11/CSIsurvey2008.pdf.

[10] D. Monahan, “Security awareness training:It’s not just for compliance,” Enterprise Management Associates (EMA), Boulder, CO, USA, April 2014.

[11] ENISA, “The new users' guide: How to raise information security awareness,” *ENISA*, https://www.enisa.europa.eu/publications/archive/copy_of_new-users-guide/at_download/fullReport, November 2010.

[12] C. Manifavas et al., “DSAPE: Dynamic security awareness program evaluation,” in *16th International Conference on Human-Computer Interaction*, Crete, Greece, 2014.

[13] J. Webster & R.T. Watson, “Analyzing the past to prepare for the future: Writing a literature review,” *MIS Quarterly*, vol. 26 , no. 2, pp. xiii-xxiii, 2002.

[14] R.C. Dodge & A.J. Ferguson, “Using phishing for user email security awareness,” in *Security and Privacy in Dynamic Environments, Proceedings of the IFIP TC-11 21st International Information Security Conference (SEC 2006)*, Karlstad, Sweden, 22-24 May 2006.

[15] H.A. Kruger & W.D. Kearney, "A prototype for assessing information security awareness," *Computers & Security*, vol. 25, no. 4, pp. 289-296, 2006.

[16] H.A. Kruger, L.Drevin, and T.Steyn, "A framework for evaluating ICT security awareness," in *ISSA 2006 from Insight to Foresight Conference*, Sandton, South Africa, 5-7 July 2006, .

[17] M. Eminağaoğlu, E.Uçar, and Ş. Eren, "The positive outcomes of information security awareness training in companies: A case study," *Information Security technical Report*, vol. 14, no. 2009, pp. 223-229, 2009.

[18] S. Talib, N.L. Clarke, and S.M. Furnell, "An analysis of information security awareness within home and work environments," in *International Conference on Availability, Reliability and Security*, Krakow, Poland, 2010.

[19] H. Kruger, L. Drevin, T. Steyn, "A vocabulary test to assess information security awareness," *Information Management & Computer Security*, vol. 18 , no. 5, pp. 316-327, 2010.

[20] E. Albrechtsen & J. Hovden, "Improving information security awareness and behaviour through dialogue, participation and collective reflection. An intervention study," *Computer & Security*, vol. 2 9, p. 4 3 2 – 4 4 5, 2010.

[21] B. Khan et al.., "Effectiveness of information security awareness methods based on psychological theories," *African Journal of Business Management* , vol. 5, no. 26, pp. 10862-10868, 2011.

[22] M. Wolf, D.A: Haworth, and L. Pietron, "Measuring an information security awareness program," *Review of Business Information System*, vol. 15, no. 3, pp. 9-22, 2011.

[23] A.R. Ahlan & M. Lubis, "Information security awareness in university: Maintaining learnability, performance and adaptability through roles of responsibility," in *7th International Conference on Information Assurance and Security (IAS)*, Melaka, Malaysia, 5-8 Dec. 2011.

[24] K. Rantos, K. Fysarakis, and H. Manifavas, "How effective is your security awareness program? An evaluation methodology," *Information Security Journal: A Global Perspective*, vol. 21, pp. 328-345, 2012.

[25] M. Wilson & J. Hash, "Building an information technology security awareness and training program," *NIST*, October 2003, https://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-50.pdf.

[26] A. Tsohou et al., "Analyzing trajectories of information security awareness," *Information Technology & People*, vol. 25, no. 3, pp. 327-352, 2012.

[27] S. Bauer, E.W. N. Bernroider, and K. Chudzikowski, "End user information security awareness programs for improving information security in banking organizations: Preliminary results from an exploratory study," in *AIS SIGSEC Workshop on Information Security & Privacy (WISP2013)*, Milano, Italy.

[28] T. Gundu & S.V. Flowerday, "Ignorance to awareness: Towards an information security awareness process," *South African Institute of Electrical Engineers,* vol. 104, no. 2, pp. 69-79, June 2013.

[29] W.A. Labuschagne & M.M. Eloff, "The effectiveness of online gaming as part of a security awareness program," in *13th European Conference on Cyber Warfare and Security*, Piraeus, Greece, 3-4 July 2014.

[30] T T. Velki, K. Solic, and H. Ocevcic, "Development of user's information security awareness questionnaire (UISAQ)," in *MIPRO*, Opatija, Croatia, 26-30 May 2014.

[31] A.N.W Prah, A.A. Otchere, and K.E. Opan, "The perceived effectiveness of information security awareness," *Information and Knowledge Management,* vol. 6, no. 7, pp. 62-73, 2016.

[32] M.C. Scholl, B. Leiner, and F. Fuhrmann, "Blind spot: do you know the effectivenes of your information security awareness raising program?," *Systemics, Cybernetics and Informatics,* vol. 15, no. 4, pp. 58-62, 2017.

[33] A. Carella, M. Kotsoev, and T.M. Truta, "Impact of security awareness training on phishing click-through rates," in *IEEE International Conference on Big Data*, Boston, MA, USA, 11-14 December 2017.

[34] D.W.H. Wahyudiwan, Y.G. Sucahyo, and A. Gandhi, "Information security awareness level measurement for employee: Case study at Ministry of Research, Technology, and Higher Education," in *3rd International Conference on Science in Information Technology*, Bandung, Indonesia, 25-26 October 2017.

[35] A. A. Al Shamsi, "Effectiveness of cyber security awareness program for young children: A case study in UAE," *International Journal of Information Technology and Language Studies,* vol. 3, no. 2, pp. 8-29, 2019.

[36] T. Gundu, S. Flowerday, and K. Renaud, "Deliver security awareness training, then repeat: { Deliver; Measure Efficacy}," in *Conference on Information Communications Technology and Society (ICTAS)*, Durban, South Africa, 6-8 March 2019.

[37] E. Ikhalisa et al., "Online social network security awareness: mass interpersonal persuasion using a Facebook app," *Information Technology & People,* vol. 32, no. 5, pp. 1276-1300, 2019.

[38] K.F. Tschakert & S. Ngamsuriyaroj, "Effectiveness of and user preferences for security awareness training methodologies," *Heliyon,* vol. 5, no. 6, 2019.

[39] J. Haney &. W. Lutters, "Security awareness training for the workforce: Moving beyond "check-the-box" compliance," *Computer* , vol. 53, no. 10, pp. 91 - 95, Oct. 2020.

[40] X. Shen et al., "Intrusive and non-intrusive evaluation of ambient displays," in *1st International Workshop on Ambient Information Systems, Colocated at Pervasive*, Toronto, Canada, 13 May 2007.

[41] I. Cereric, J. Looney, and M. de Greef, "Indicators for evaluation of awareness and fundraising for low literacy in Europe," ELINET- European Literacy Policy Network, Brussels, Belgium, 2014.

[42] M. Kajzer et al., “An exploratory investigation of message person congruence in information security awareness campaigns,” *Computer & Security*, vol. 43, no. 2014, pp. 64-76, 2014.