Article

Behaviour coding approach for assessing pitfalls in a questionnaire instrument towards assessing healthcare security practice

Prosper Kandabongee Yeng [1,†,‡], Muhammad Ali Fauzi [1,*] and Bian Yang [1,*]

[1] Norwegian University of Science and Technology; prosper.yeng;muhammad.a.fauzi; bian.yang@{ntnu.no}
* Correspondence: prosper.yeng@ntnu.no
† Current address: NTNU, Teknologivegen 22, 2815 Gjøvik
‡ These authors contributed equally to this work.

Received: date; Accepted: date; Published: date

Abstract: Questionnaires are useful instruments for gathering responses to specific factual questions. However, the problems of questionnaire responses impede the effective use of the questionnaire. Some of these problems including non-responses, non-completion, issues of judgement, and social desirability such as information that the respondent is not willing to disclose need to be dealt with. The situation is more compounded in a scenario where an information security practice study in a typical hospital consists of broad categories of respondents and the survey findings are to be relied on to actually address information security compliance issues. This study, therefore, shares "pitfalls" to watch when preparing a questionnaire to measure the information security practice level in a hospital which is characterized by different respondents with varying domain knowledge such as knowledge in information security, information communication technology, and that of the domain knowledge of healthcare. A synergy of a conventional pretesting method and behaviour coding were therefore used to pretest the questionnaire. Questionnaire problems including a lack of understanding of the healthcare information systems' structure of all hospitals, unclear questions, the insignificant difference between questions, problematic questions, inadequate questions, and complex terms were among the identified pitfalls to watch. Out of a total of 118 questionnaire items that were used in the pretesting, a total of 50 questionnaire items (representing 42%) were identified to have problems after the pretesting was conducted with a total of 36 respondents in behavior coding and 21 respondents in conventional pretesting.

Keywords: Security practice; Healthcare; Questionnaire design; Questionnaire pretesting

1. Introduction

Digitalization has largely been adopted across the globe in efforts to enhance efficiencies in healthcare. For instance, telemedicine systems, Electronic Health Record (EHR) management systems, implantable medical devices (such as pacemakers), and many others have tremendously transformed the effectiveness of our healthcare provisions and improved upon the quality of life of patients as well [1]. However, digitalization in healthcare has also increased the attack surface [2]. Traditionally, technological countermeasures (such as firewalls, antivirus, intrusion detection, prevention systems, etc.) have been the default security solutions and have matured over time. These countermeasures are capable of preventing the circumvention of cyber-criminals to illegitimately gain access to healthcare systems. As a result, the hackers have shifted their mode of nefarious operations. They gain unauthorized access into healthcare systems through the healthcare worker [3,4] who is believed to be the easiest target and the weakest link in the security chain [3,4].

In 2017, the healthcare system of the United Kingdom was heavily impacted by the Wannacry ransomware and this affected critical care [6,7]. The ransomware spread to about 150 countries and affected about 230,000 computers in different sectors. Following that, in 2018, about 3 million healthcare records were compromised in Norway [8,9] of which an insider aid was underscored. According to

 doi:10.20944/preprints202212.0369.v1

HealthCare IT News, there was another phishing attack that led to a breach of 38,000 patient records in Portland, Oregon-based Legacy Health in the United States in 2019. Personal data such as patients' email accounts, demographic information, dates of birth, health insurance data, billing details, medical data, Social Security numbers, and driver's licenses were stolen. Healthcare data breaches continue to increase sharply, with the passage of time. According to Verizon, globally, about 5 million healthcare records were compromised in 2017, followed by 15 million records in 2018 and 25 million records in the middle of 2019[13]. This is threatening the quality of healthcare [14].

In efforts towards providing solutions, a project called Healthcare security practice analysis modeling and incentivization (HSPAMI) project has been initiated at the Norwegian University of Science and Technology (NTNU) [8,10]. The general objective of this study is to understand the challenges often faced by healthcare workers in their effort to comply with security controls while executing their duties. The result is intended to guide in determining incentivization methods and developing better security controls towards strengthening the "human firewall" in order to enhance the security within healthcare systems. The "human firewall" is the fortification of the security-conscious care behavior of the healthcare staff towards preventing unauthorized access to the systems.

The security practice involves information security measures being adopted towards complying with the confidentiality, integrity, and availability (CIA) of information systems and other assets. Security practice is relative to various healthcare groups of workers. For instance, end-users such as doctors, nurses, paramedical staff, pharmacies, and health administrators indulged in various security practices including password management, social media use, internet use, mobile system usage, and email use. Additionally, information security officers in healthcare are also responsible to perform various activities including backups, threat modeling, and systems security updates.

In efforts to improve on the security practice, a survey is being carried out to determine the effect of the psychological, social, and cultural influence of the healthcare staff's characteristics on various security practices. Additionally, EHR logs of the healthcare staff are to be analyzed to determine deviations of healthcare staff security practices. Healthcare staffs are often attacked by hackers through social engineering techniques. Therefore, simulated studies will be conducted to determine the status of the healthcare staff in terms of their ability to resist social engineering tricks [10,18].

Questionnaire instruments, interviews, and observing how healthcare staff comply with the security measures are some of the approaches that are to be adopted in the survey project. However, there are unique challenges in modeling and analyzing information security practices using a questionnaire instrument and interviews within the healthcare sector. This is because the healthcare sector is characterized by various categories of study participants. For instance, in the information technology (IT) department in a healthcare setting, there could be IT offers in which some are responsible for information security management, application development, IT project management, and network administration. Additionally, there are also healthcare professionals whose primary role is to offer healthcare services with the support of the IT systems. These healthcare staffs include doctors, nurses, pharmacies, laboratory personnel, and radiology personnel. Within the healthcare personnel, some of them have professional knowledge in IT while others have acquired basic IT skills just to enable them to do their work. Furthermore, the healthcare sector consists of other supporting staff such as the health administrators, human resource personnel, finance, and accounting personnel, and they equally rely on the IT systems in the discharge of their duties.

Due to IT knowledge gaps across these categories of healthcare staff, developing a questionnaire to effectively measure the security practice while reducing response burden, require a multifaceted design and pretesting strategies to balance the understanding of the constructs and all questionnaire items among the respondents. For instance, a nurse who studied IT security at the bachelor or master level will have a higher level of understanding of some terminologies in an information security questionnaire items than their colleague nurse who has been trained on how to only use the IT systems to do their work. Besides, the way the questionnaire items will be structured to have a common denominator that can be able to efficiently examine the security practice of the multifaceted categories

 doi:10.20944/preprints202212.0369.v1

of respondents in their respective healthcare roles should also become a big concern. Moreover, in a comparative study involving different countries, it is also important to identify the resources required to enable respondents to successfully answer the questionnaire. Various hospitals in different countries may have different IT setups in their hospitals. In scenarios whereby in one country, the hospitals only operate on local area network while in some other countries, the hospitals operate on a wide area network with the support of the internet, how will the questionnaire items be scaled to have a common denominator for the assessment of the security practice of healthcare staff across such variances? Even in the common use of online administration of questionnaires in this current availability of IT solutions, which online questionnaire administering system is secure and reliable for collecting sensitive responses from healthcare staffs?.

The specific objective of this study was to survey for various questionnaire designs and pretesting methods. The methods were then assessed for their suitability for designing and pretesting questionnaires for healthcare security practice. The appropriate methods were then used to design and pretest our questionnaire for collecting data on security challenges often faced by healthcare professionals in their effort to comply with security requirements. The findings and experience in the design and pretesting of these questionnaires have also been shared in this paper.

Section 2 presents various questionnaire design methods and how they were used in pretesting our security practice questionnaire. This is followed by an overview of pretesting methods in section 3. Section 4 showed how the design and pretesting of the questionnaire were carried out in this study. Section 4 and section 6 respectively presented the pretesting findings and discussions.

2. Questionnaire design Methods

A questionnaire is the most widely used instrument for gathering data on knowledge, attitude, opinion, behavior, facts, etc., concerning people, objects, and events [66,67]. The questionnaire tool is usually developed to collect information from the respondents. Various questionnaire design methods such as the rational, the prototypical, the facet design, the construct, the internal, and the external methods are usually adopted [15] [18] as shown in 1.Additionally, there are four stages involved in questionnaire development: These are concept analysis, item reproduction, scale construction, and evaluation. Each of the questionnaire design methods and how they are applied to these stages is shown in table 2.

The rational method [15] relies solely on the knowledge of the questionnaire developers or expert's judgment to place the questionnaire items [16,17]. So the experts mostly consider in their perspective, what is rational or reasonable based on their knowledge in the problem domain. The experts' knowledge in placing the questionnaire item is usually guided by the face validity, to determine if a questionnaire looks valid to measure the desired construct of interest from the opinion of the expert or developer, respondents, and other observers who are technically untrained.

The prototypical method is based on a theory from cognitive science called prototype theory. It is about the presentation of categories such that the variation of members of a category degrades away from the center [15,17,64,65]. So the objects or members of a class are mostly related to that class. Literally, the prototypical approach is like clustering of the questionnaire items where the items members of a category vary in characteristics. So members which are more related (more prototypical) to a category or a class are easier to be categorized. In the test construction stage of a questionnaire, each construct is represented by a set of acts or behavior that are more prototypical related or central to that construct. The prototypical approach concentrates on placing items under the constructs that are more related in characteristics and this increases the cognitive processing of respondents thereby, increasing the quality of the questionnaire instrument. The construction is usually guided by the informal knowledge and experience of the respondents [19,20]. In the item production stage of the questionnaire, members of the target population are given the option to nominate experts with vast knowledge in the construct, to outline behaviors that relate to this construct. The editing of the expects'

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 December 2022 doi:10.20944/preprints202212.0369.v1

constructs by the questionnaire developers is highly limited in order to preserve the incorporation of the expert's knowledge of the questionnaire items.

The facet design method [15,18] uses the principle of content validity. It adopts a comprehensive and systematic approach to analyze and specify the construct towards ensuring that the items in the questionnaire have significantly represented the construct. The process begins with a comprehensive catalog of the construct domain and categorizes it into various sections referred to as the facet [21,22]. Each facet is further expanded into facet elements. The facet design method is also a hypothesis testing method with empirical tests much like the construct method. But this method differs from the construct method in the sense that it does not primarily depend on a formal theory of the construct. It only requires formal knowledge of the construct and goes through four steps in the concept analysis phase. These include taking an inventory of the behavioral features and their underlying processes that are relevant to the construct. An example is a security practice that has psychological, social, cultural, and demographic factors. To ensure content validity, all these aspects need to be represented in a questionnaire in a typical study on security practice as shown in the independent variables of Figure A1.

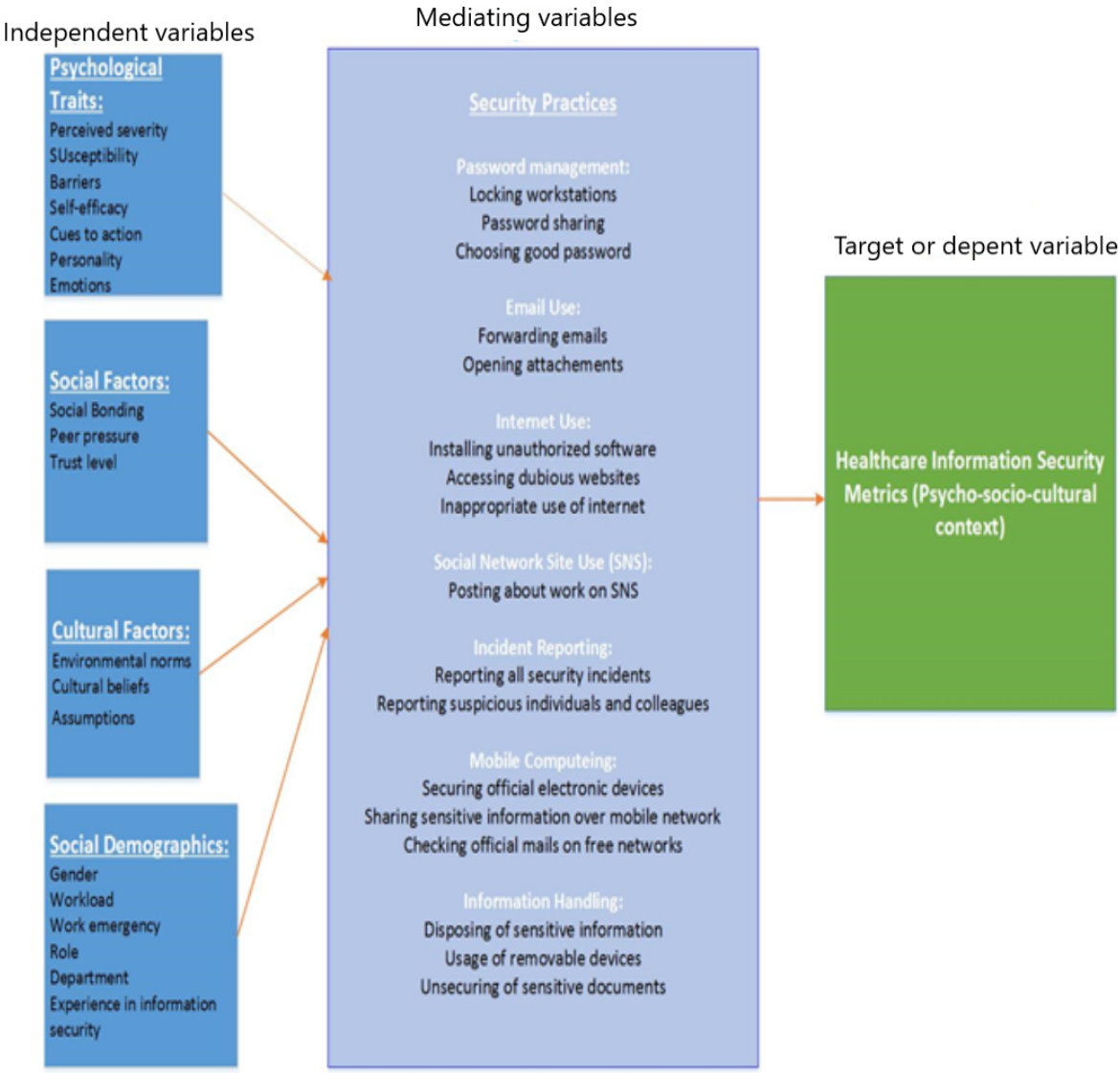


Figure 1. Nomological network with independent, mediating and dependent variables.

Secondly, the facets are designed to be mutually exclusive by expanding the inventory. For instance, the psychological facet of the construct may be expanded to include psychological attributes

 doi:10.20944/preprints202212.0369.v1

such as perceived severity, perceived vulnerability, and perceived efficacy. In the third step, the questionnaire items of each facet can then be determined to cover the content domain. Finally, the various items in each facet are then combined to form the formal structure [23].

The construct method [15,18,24,25] is primarily derived from a theoretical approach in which the constructs are developed from the theoretical concept of the research domain. Hypotheses about the questionnaire are then developed and assessed empirically. if the questionnaire items or scales are found to deviate from the construct theory, new construction of the items and scales are formed by revising the questionnaire. The concept analysis of the construct method relies on the guidance of the construct theory, in which a nomological network with relevant variables (independent, mediating, and dependent) and their respective relationships among the variables are often used. The related variables usually have correlation with the construct in the study [15,18,24,25].

In the internal method [15,18,26–28], constructs cannot be specified in advanced. Constructs are formed or derived from empirical relations between questionnaire items. The co-variance among a set of questionnaire items is due to a common factor and that is considered as the underlying construct. The internal method is largely used to improve upon the existing questionnaire instrument. It can also be used to create a new questionnaire from a pull of questionnaire items in the same domain. Since the construct is not pre-specified in the internal method and constructs are specified, the concept analysis stage of a questionnaire is often skipped. So the question construction often begins with the item's production.

The external method[15,18,29] fundamentally relies on the principle that individual responses to a questionnaire item represent their respective behavior that may be related to many non-test behaviors. The statistical relationship between the items responded and the specified behavior is therefore considered to be more informative. The construction does not require theories or concept specification but the concepts are obtained through the predicted or determined non-test behavior or criterion and not through psychological theories. A collection of heterogeneous items that are related to the criterion are gathered at the item production stage. The item pool consists of heterogeneous items of which the diverse items are very important. The many different aspects of the items in the items pools significantly contribute to the production of non-test behavior. The strength of the relationship between items and the study area (criterion) is much focused on the scale construction stage.

A summary of the methods and their objectives are shown in table 1.

Table 1. Questionnaire Design Methods and Objectives.

Method	**Objective**
Rational	Most applicable when little formal knowledge of the concept is known or when the concept has not been explored into details
Prototypical	It is aimed towards increasing the cognitive process of respondents. Therefore, more prototypical items would be better understood
Facet	The intention of this method is to ensure that the questionnaire items comprehensively represent the constructs. This is achieved by systematically outlining the constructs with the guidance of content validity
Construct	The construct method aimed towards designing questionnaire items to meet the construct theory. This is done by relying on formal knowledge to generate a hypothesis of the items for empirical assessment
Internal	The method is intended for improving upon existing questionnaires or developing new questionnaire items from existing questionnaires that are related. This is done by observing the common factor or co-variance among the items
External	This method aimed towards a generalization of the questionnaire with an external criterion by gathering and analyzing for the strong relation of heterogeneous items that are related to the construct and the external criterion

The questionnaire design methods are usually applied at various stages of the design as shown in Table 2. The concept analysis is the initial stage of the questionnaire construction. So the theoretical framework is identified together with outlining and defining the constructs. Primarily, the concept

 doi:10.20944/preprints202212.0369.v1

stage includes the specification of the purpose, objectives, research questions, and hypotheses [66]. At this stage, the respondents' background needs to be understood and a clear understanding of the problem needs to also be explored through literature reviews.

Table 2. Questionnaire design methods and their applications in the various design stages.

Design Methods	Questionnaire design Stages			
	Concept Analysis	**Item Production**	**Scale Construction**	**Validation**
Rational	The theoretical framework of the concept is usually originated from the ideas of the questionnaire developer; Concepts are obtained in a typologies syndrome or global descriptions.	Uses intuitive or informal criterion in the item production; Items are produced using typologies, syndrome and global description.	Scale construction is based on the developer's judgement	Items are assessed based on face validity. Estimates of the validity and reliability are also evaluated.
Prototypical	-	Item production is based on act-nomination. Thus persons with extreme knowledge of the construct are selected to outline behaviours that illustrates the construct. Developers of the questionnaire are not required to make significant changes.	Prototypical ratings are used to select the questionnaire items	Peer rating procedure is used in evaluation [30,31].

 doi:10.20944/preprints202212.0369.v1

Table 2 – continued from previous page

Facet	The following steps are followed: An inventory of the behavioural factors and underlying processes of the construct are taken; The facet are defined via elaboration of the inventory; The facet elements are then determined to be mutually exclusive; The item in each facet are then combined to from the final study.	The total number of items needed, depends on the size of the facet design per facet. Content is then created for each facet element.	Items with high correlation to the intended scale are optimised for used	Network confirmatory factor analysis, reliability analysis and different item function
Construct	This is guided by the construct theory, expressed in a nomological network, using relevant variables to indicate their relationship, as shown in Fig 1.,	The operational definition of the construct in the study is used to generate the items. Experts and potential respondents can judge the items	Scale construction is based on the content saturation ie the convergence and discriminant validity of the of the items. Items with high correlation to the intended scale are optimized for use	-
Internal	-	Items are selected to be relevant to the content domain with a degree of content homogeneity. Existing set of items are often selected similar content domain.	The focus is on the homogeneity of items. Homogeneous items are then interpreted post-hoc with their respective scaling. Non-homogeneous items are excluded	Stability of identified item co-variance structure is mostly assessed using cross-validation and confirmatory test
External	-	Collection of heterogeneous items that are relevant to the criterion is obtained[32]	The strength of the relationship between and criterion are used. Items with higher correlation with the criterion but low correlation among them are optimal	Cross validation is used for item-criterion and scale-criterion assessment. Test-retest is also used.

The item production stage, which is also referred to as the questionnaire conceptualization stage [66], involves gathering of the questions or obtaining a pool of the questionnaire items based on the initial specification in the concept analysis stage. In this phase, the items can also be reviewed by experts or potential respondents. The content of the questionnaire (eg. the questions) is developed by

 doi:10.20944/preprints202212.0369.v1

specifying whether the questionnaire is intended to measure either knowledge, attitudes, perceptions, opinions, recalling facts, behavior change, etc. Variables including independent, mediating, and dependent variables are also specified. Essentially, pretesting of the questionnaire can be conducted at this stage and the findings can be used to subsequently update or revise the questionnaire items. Meanwhile, in the scale construction stage, possible answers to each questionnaire item are selected for the scales while relying on options that increase the psychometrics (mental capacity and processing) of the study. Having a good understanding of the collected data and the suitability of the data analysis method is very necessary. Therefore a preliminary statistical analysis method should be tested to ensure that the results will meet the needs of the study objective.

The final step involves evaluation of the questionnaire which consists of establishing validity and reliability. Validity here refers to the measure of what the questionnaire claims to be measuring [66,67]. Validity can be reduced if a higher proportion of errors are systematically built on the questionnaire instruments. The types of validity include content, construct, criterion, face, internal, and external validity.

Internal validity is concerned with the effectiveness of the results within the scope of the study. It has to do with causality, i.e. the assignment of causes to outcomes. External validity is the extent to which the study can be applied or generalized beyond the current study scope. A construct is a concept being measured by a test. So construct validity is a test to assess if the study is measuring the construct it was designed to measure[68,69]. A higher number of different measures increases the construct validity on the condition that the measurements are all measuring the same concept or construct in the study. Construct validity is mostly assessed by comparing the study to other tests that measure similar qualities to determine the correlation between the two studies. Content validity is when a study measures the knowledge of the content domain of which it was designed to measure. For instance, if a study was designed to measure information security practice in healthcare, the content validity assesses if the study was indeed measuring information security practice and not measuring something else outside security practice such as computer use. Furthermore, if the study area is about information security practice, the content validity criteria are met when the questionnaire items have comprehensively covered all the knowledge areas of information security practice [68]. Face validity involves assessment of the questionnaire to determine if the test appears valid to the respondents, the administrator of the questionnaire [68]. Criteria validity is a measure that seeks to determine how well a test can predict the outcome of another measure[68]. An instance includes how well can the results of the information security questionnaire be used to sufficiently provide incentivization methods to improve upon the security-conscious care behavior of users. Another example is how well a security practice questionnaire results will be used to correct the challenges often faced by healthcare staff as they go about with their usual duty [68,69]. A panel of experts and field tests can be used to assess the validity.

3. Pretesting methods

Pretesting aims to identify ambiguous questions or wording, unclear instructions, and other issues with the questionnaire prior to actual use [33,41]. It helps to also provide information concerning reliability and validity by identifying potential problems before the actual use of the instrument. So the pretesting of the survey instrument should be fully described. The type of pretesting includes conventional pretesting, cognitive interviews, behavior coding, response latency, vignette analysis and, formal response debriefing. Other types of pretesting include focus group, group discussion, experiment and statistical modeling [33–36,41] as summarized in Table 5.

3.1. Conventional pretesting

Conventional pretesting of the questionnaire is based on the assumption that the problems of the questionnaire are identified based on issues raised by respondents. Respondents may refuse to answer such items or indicate their lack of knowledge on the item. In conventional pretesting, there

 doi:10.20944/preprints202212.0369.v1

are no exact tests but the testing is based on subjective views of respondents which are translated or transcribed by the fieldworkers or interviewers[33–35] and this can make it difficult to understand the full problems raised by the respondents. Moreover, in the debriefings of the interviews in conventional pretesting, there is no scale to illustrate the severity of the described problems. Therefore, researchers only rely on intuition and experience to judge the seriousness of the problems and decide on what to address [33]. Additionally, respondents may be hesitant, uncomfortable, or showed visible resistance in answering the questions which interviewers can note down during the debriefing session [33] but this will require the presence of the researcher during the response process. This requirement, therefore, drifts into a cognitive interview.

3.2. Cognitive interviews

The cognitive interview involves drawing out the view or depicting the processes in answering questions. This is towards analyzing to extract the exact problems associated with the questions [33–39,41]. The intention is to reveal the thought processes in interpreting and answering a question.

The methods employed in cognitive interviewing include verbal probing (concurrent and retrospective), observing the respondent's behavior and think-aloud, or read-aloud while the respondent answers the questions in the questionnaire instrument [38,39].

The probing involves asking the respondents to interpret the questions and to provide an understanding of the meanings of words used in the questions. The respondents also explain their responses and pinpoint areas of the questionnaire that has difficulties in understanding, interpretation, or completion [40,41,44]. The fundamental goal is to clarify the understanding of respondents' on the questions.

Observational aspect in cognitive interviews includes observing respondents skipping questions, flipping a page back and forth, placing answers in the wrong fields of the questionnaire form. The observation also includes changes in appearance such as frowning and hesitation. In this regard, the respondent can be directly questioned about the difficulties they are facing with each question on the questionnaire.

In the think-aloud method, the researcher aims to understand the cognitive processes used by the respondents in answering the questions. So the researcher encourages the respondents to speak out their thoughts while answering the questionnaire instrument [41–43].

In a related study, Cantril and Fred used one or two written probes to assess the understanding of survey questions from respondents. Similarly, Balson used none directional probing to explore seven questions which were previously responded through ordinary interview. Ordinary interviews aim to obtain just the responses to the questions [33]. Retrospective probing was used by the interviewers to extract the cognitive process employed by respondents to arrive at their answers in their previously responded questioned which ordinary interview was used. Belson's hypothesis that respondents could reconstruct their thought process from their previously answered interview. Belson's study [48] did not have much impact on pretesting practice due to the adoption of labor-intensive in the study. Subsequently, the cognitive interview became popular after think-aloud was used to answer survey questions about past events [33,47]. Cognitive laboratories were also created towards understanding responses to the survey questionnaire [33].

Cognitive methods that involve verbal reporting of information in short-term memory are likely to be veridical if the verbalization does not involve further explanation that has the potential to alter the thought process [33,46]. But the general use of verbal report methods in cognitive processes for answering survey questionnaires is difficult to justify particularly for tests that fail to meet the conditions for valid verbal reports such as term comprehension. Although the social interaction between respondents and interviewer in a cognitive interview may violate key assumptions [33,46], a certain kind of verbal probing poses difficulties for respondents. Respondents do not have any difficulties with re-orienting probes (asking for an answer). However, respondents have difficulties with elaborating probes (asking for further information). Also, a cognitive interview is much directed

 doi:10.20944/preprints202212.0369.v1

by the interviewer and as a result, the outcome can be biased from the interviewer interference and this could fail to support the inquiry.

3.3. Behaviour coding

Behavior coding involves observing subsets of the respondents' and interviewer's verbal behavior by monitoring interviews or reviewing recorded interviews of their question asking and answering interactions [33,49]. Questions found to be scoring high frequency (15% or above of the respondents) in a certain behavior (such as the respondent needed clarification or the interviewer failed to read the question verbatim) are noted to need repairs. Aside from the frequency count in certain behavior, the sequence of a question's behavior coding is coded as either paradigmatic, problematic, or inadequate [33,50] s detailed in Table 3.

Paradigmatic involves a question coding where the interviewer correctly reads the question, the respondents select one of the offered alternatives, then the interviewer codes the answer as being correct and this is also referred to as adequate answering as shown in Table 3. Problematic codes mean that the sequence was non-paradigmatic but the issue was resolved. For instance, the respondent asked for clarification, and then selected one of the offered alternatives. Inadequate coding is also non-paradigmatic but the identified problem remains unresolved. Questions with a high proportion of non-paradigmatic sequences such as inadequate answer, don't know, or refused to answer are also tagged as needing repairs.

Table 3. Behaviour codes [41]

Behaviour Type	**Meaning**
Reader/Interviewer behavior Codes	
Exact (E)	Interviewer exactly read the questionnaire item
Slightly Changed (S)	Interviewer made minor wording change when reading the question
Major Changed (M)	Interviewer made significant wording change when reading the question
Respondent behavior	
(1) Interruption with answer	Respondent interrupted the question reading to give his or her answer
(2) Clarification	Respondent asked for clarification
(3) Adequate Answer	Respondent provided answer after question reading
(4) Qualified answer	Respondent qualified his or her answer
(5) Inadequate answer	Respondents initial answer was inadequate
(6) Don't Know	Respondent gave a "don't know" response
(7) Refusal to answer	Respondent refused to answer the question

3.4. Response latency

Response latency is another testing method which has to do with the among of time respondent takes to answer a question [51]. Longer delays in response latency signify uncertainty in the response of the respondent's answer and this inadvertently evaluates questions by themselves. There was an association in a multivariate analysis in which both longer response latencies and the expressions of the respondents' uncertainty about the responses. Certainly, the interpretation of the response latency is not straightforward because a long time needed may be due to careful processing instead of difficulties in interpreting the questions. Aside, the method appears significantly effective for its adoption and can be combined with other testing methods to enhance effectiveness.

 doi:10.20944/preprints202212.0369.v1

3.5. Respondent debriefing

Respondent debriefing involves the incorporation of follow-up questions in a standard interview [33,52]. The respondent is duly informed about the purpose of their response inquiry [33]. In pretesting questionnaire with respondent debriefing provides knowledge on the meaning of the questions and the reactions of the respondent to the question [33].

3.6. Vignettes

Vignettes[33] involves hypothetical scenarios that respondents assess and this may be conducted in either undeclared or participating pretesting methods. Vignettes approach is more suitable for exploring how people think about concepts; test the consistency of the interpretations of the respondents, tests the consistency of the interpretations of the concepts of a respondent, assess the dimensions associated with a concept, and assess other questions word problems.

3.7. Experimental methods

Supplementary methods to conventional pretesting and cognitive interviews can both identify various problems of a questionnaire instrument. This leads to a revision and redesign of the questionnaire to address the identified problems. But how will it be known if the revisions are an improvement to the old questionnaire? An experimental comparison of the original and the updated version is used. One way of the experimental comparison is to compare the original and the updated versions of the questionnaire, using the initial test methods which were used in the pretesting of the original version [33] to identify the problem. For instance, if conventional pretesting was supplemented with response latency to identify problems with an item, that item and its revised version can be tested again with the conventional pretesting and any supplemented method (such as response latency) which was used. This is to ascertain if the revised version indeed had fewer problems. Furthermore, the original and the updated version of the questionnaire can be tested to determine if there is a difference in the survey estimates.

3.8. Focus group

When the construct and objective of a questionnaire are set, a focus group is then formed to meet. The group then assesses respondents. Areas of the questionnaire in which required are listed. The understanding of tasks such as keywords, concepts, recalls of respondents, question-wording, etc are assessed [71].

3.9. Expert reviews

Expert review is among the commonest methods in questionnaire pretesting and it involves individual or group sessions of experts who completely rely on the expertise to identify problems in questionnaire items[44,77,78]. The experts usually follow a structured appraisal system or rely on their own informal judgment to identify the questionnaire problems.

3.10. Statistical modeling

Statistical models such as latent class analysis(LCA) [33,53] and item response theory (IRT) [33,54] can be used to estimate the errors in questionnaire instrument. LCA is similar to clustering that identifies hidden sub-groups in a population in which the subgroups are unique to each other but have similar members or objects. When the questions have been answered by the same respondents two or more times, LCA can be used to analyze and estimate the errors associated with the questions as described in [53]. IRT determines a way of testing the understanding of respondents by collecting answers from each respondent and indicating whether their respective answers were correct or not. The results are then organized to indicate each respondent and the status of their respective

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 December 2022 doi:10.20944/preprints202212.0369.v1

responses (correct answer or wrong answer). The data is then analyzed towards improving upon the questionnaire [70].

Table 4. Pretesting techniques and their objectives

Method	Objective
Conventional pretesting	To obtain subjective evaluation of the questionnaire from the answered results of respondents
Cognitive interview	Focuses on understanding the process of answering questionnaire in order to evaluate the understanding of the objective of the questionnaire
Behaviour coding	A third person observes the interaction between the interviewer and the participant to identify signs of misunderstanding or difficulties based on the occupancy frequencies towards improving upon the identified items [41,55]
Debriefing	Much like probes,the objective is to determine comprehensiveness and problems associated with the questionnaire item[41,55]
Response latency	To identify difficult or complex questions through delays of respondents which are correlated with their uncertainties
Vignettes	Assesses how respondents understand concepts in different wording scenarios, questions and wording problems in relation to the goal of the study [33,41]
Experimental	Aim to determine errors in questionnaire by repeatedly asking the same respondents,to answer a questionnaire at a given number of times and the results analysed
Focus groups	Assesses the way respondents understand keywords, terms and phrases
Expert reviews	Designed questionnaire are reviewed by experts for possible problems [77]

4. Our Approach

The aim of this paper was to effectively design and pretest a questionnaire to improve the response rate of the questionnaire. Literature concerning questionnaire design methods and pretesting were hence surveyed in Google scholar, IEEE Explore, PUBMED, Science Direct, Elsevier, and SCOPUS. The keywords which were used in the search include Questionnaire design, Questionnaire pretesting, Information security, and healthcare. Boolean functions of AND, OR, and NOT were used in combining the search terms for a better search strategy.

Each of the questionnaire designs and the pretesting methods was assessed and considered for the design and pretesting of the healthcare security practice. The assessment of the design methods was related to the objective and scope of the study of healthcare security practice questionnaire objective in relation to the psychological, social, and cultural context. With respect to the questionnaire design methods as shown in Table 1 each of the objectives of the design methods was accessed with that of the HSPAMI questionnaire objectives. For instance, the prototypical method presents categories of items with similar members and with the objective to increase the cognitive aspect of the questionnaire as shown in Table 1. In that light, the independent variables of the nomological network (see in Figure 1) are categorized into various aspects including psychological traits, social factors cultural factors, and social demographics. Based on that, the prototypical method is considered suitable for designing the questionnaire for healthcare security practice in that aspect.

Furthermore, our pretesting was conducted by using a conventional pretesting approach and behavior coding having analyzed their pros and cons.

 doi:10.20944/preprints202212.0369.v1

Conventional pretesting is a faster process of questionnaire evaluation. So this method was used to have a quick idea of whether our questionnaire had issues. However, conventional pretesting is unable to identify all questionnaire problems since respondents can wrongly answer the test questions based on their wrong interpretation of the questions. Conventional pretesting alone is not able to identify many pretesting problems because some kind of problems cannot be observed from the responses of the respondents, especially if the respondent themselves are not aware of the problems [33]. In a scenario where the respondent is not aware of the intention of a closed question, the person may miss-interprets the understanding of a closed question without leaving any sign of such behavior. Additionally, in a conventional pretesting approach, the fieldworker or interviewer is exposed to the intention of the pretesting exercise. Whereas the intention of the pretesting may remain undisclosed to the respondents. So the respondents are mostly not able to ask relevant questions that meet the intent of the question, unlike the interviewer who has the privilege of knowing that [33,36]. Fortunately, these gaps can be covered by complementing with any of other pretesting methods such as cognitive interview, behavior coding, debriefing, response latency, vignettes, experimental method, focus group, or expert review methods[33].

Vignettes involve multiple testing which enables an interviewer to have a comprehensive view of the question problem. It also combines the strength of both internal and external validity but the vignette is considered artificial because it adopts hypothetical behavior. However, hypothetical behavior in an experiment may not reflect the behavior of one's real-life [62]. The experimental approach has also been considered to be easy to implement but it greatly relies on the assumed module and lacks effective evaluation methods [53].In a focus group, the thought process can be stimulated by the comments of others. The members' reaction and their comments can provide valuable ideas to help in revising the questionnaire items [63]. However, the results are difficult to implement and it also takes a lot of time to organize and process the results and in the end, very few terms, topics, or issues are addressed [41].

In view of these, behavior coding was combined with the conventional pretesting in this study. Behavior coding pin-points to the problem of the questionnaire item rather than providing the causes as compared to other pretesting methods such as cognitive interviewing, vignettes, debriefing, and focus groups [83] as listed in Table 4 . Notwithstanding, behavior coding results are more reliable since the data are collected in a situation that reflects the main study. Furthermore, whereas other pretesting methods such as cognitive interviewing, may report pseudo-problems in order to just please the interviewer, respondents of behavior coding report the real problems. Moreover, behavior coding is more objective, reliable, easy to do and it provides data under realistic conditions [82]. Behavior coding is also a comparatively simpler and low-cost method[41,83].

As a result, the designed questionnaire was sent to respondents in Norway, Indonesia, and Ghana. The target group was healthcare staff who use information systems in providing healthcare services to patients in the hospitals. The pretesting methods were conducted with various types of respondents in healthcare including nurses, doctors, IT researchers in healthcare, and laboratory personnel who were working in hospitals that were not part of the main study healthcare facilities. Conventional pretesting questionnaire in English was first sent out to the target respondents in various hospitals in Ghana and Norway and having independently obtained ethical clearance from the research councils of these countries. Subsequently, invitations were sent out to healthcare workers in these countries to help in testing our questionnaire using the behavior coding method. Respondents were motivated to participate by refunding each of them an amount of GHS 10.00 for using their internet data. A synergy of these methods was used to develop and pretest the questionnaire for healthcare security practice.

Table 5. Questionnaire design methods used for designing questionnaire in healthcare security practice

Method	How it was used in HSPAMI	Example

 doi:10.20944/preprints202212.0369.v1

Table 5 – continued from previous page

Method	Reasons	Example
Prototypical	From Figure 1, both independent and mediating variables have different categories. This method was used to present questionnaire items based on all the categories, to enhance cognitive processing.	As shown in figure 1, the questionnaire items will be categorized into the various constructs of the nomological diagram such as psychological, social and cultural and demographic variables such that the item within each categories are related but differ in various categories.
The facet method	This method was adopted due to its objective to adequately represent the questionnaire items to meet the content validity in a systematic manner.	With reference to the questionnaire in appendix 1, each construct is further categorised. For instance, in the social construct, sub-construct were identified to include Formal control and informal controls of which questionnaire items were developed for each of these constructs towards enhancing the content validity of the questionnaire
The Construct	The entire psycho-socio-cultural aspect of the HSPAMI study is based on a theoretical background as shown in figure 1, where psychological, social, cultural theories have been combined to form individual traits. The construct method was adopted in the design so that the questionnaire outcome will be able to adequately assess each of the constructs	An instance is the construction of the questionnaire to consider all the constructs (independent variables) as shown in figure 1
Internal method	The method was used in developing questionnaire based on existing questionnaires, since the method focuses on empirical relations between questionnaire items and mostly used for developing questionnaires based on existing questionnaire items	From appendix 1, existing questionnaire items that are related to the various mediating variables (security practices) were selected from [72,74–76,81]

5. Pretesting results

Having assessed the identified questionnaire design and pretesting methods based on their objectives and their application as shown in Table 2 and Table 4 respectfully, a number of them were suitable for designing and pretesting a questionnaire towards analyzing healthcare security practice. This questionnaire was then pretested with conventional pretesting, and behavior coding. The findings of conventional pretesting are shown in Table 6.

A total of 21 respondents answered the questionnaire in the conventional pretesting approach and provided feedback on various issues which were categorized to include an arrangement of questions scales (19%), incomplete response options (19%) and unclear questions (19%) as shown in Table 6. A relatively high number of respondents (23.8%) opined on complicated terms in the questionnaire items.

 doi:10.20944/preprints202212.0369.v1

Table 6. Conventional pretesting findings

#	Questionnaire issues identified	Count of Respondents	%
1	Arrangement of questions scales	4	19.0
2	Incomplete response options for respondents	4	19.0
3	Understanding of information systems' structure of hospitals	2	9.5
4	Use of abbreviations which are not understood by all respondents	3	14.3
5	Questions that are unrelated to the responsibilities or task of respondents	1	4.8
6	Unclear question	4	19.0
7	Complicated terms for respondents eg multi-factor authentication	5	23.8
8	Insignificant difference between some questions	4	19.0
9	Security and privacy need of respondents in relation to the questionnaire tool	1	4.8
10	Grammatical errors and spelling mistakes	3	14.3

5.1. Behaviour coding results

A total of 36 respondents were interviewed in the behavior coding method with the questionnaire which contains a total of 118 questionnaire items in 6 different sections. The questionnaire was first revised in English based on the results from the conventional pre-testing. Since the same questionnaire was to be answered in Norwegian and in the Indonesian language, it was translated into these languages as well. The pretesting was done in Ghana, Norway, and Indonesia by respondents who had healthcare backgrounds and were not working with the hospitals in which the actual study was to be conducted. Among the 36 respondents, 10 of them were interviewed in English, 7 of them preferred Norwegian and 19 of them were interviewed in the Indonesian language. So a total of 4,248 (36 respondents * 118 questionnaire items) coded behavior were collected from the respondents and the interviewers. The results are based on the responses across the three languages (English, Norwegian, and Indonesian).

5.1.1. Respondent behavior types

Table 7 showed the codes of the respondents behaviour types and the quantification of the responses from the respondents. Out of the 4,248, 3923 responded with adequate answers (3). This represents about 92.35% of the total responses. A total of 262 responded with clarification (2) and this represented 6.17% of the total responses. None of the respondents refused to answer any of the questionnaire items as shown in Table 7.

Table 7. Proportion of respondent behaviour types

Behaviour Types Codes	Respondent behavior counts	%
1 Interruption with answer	5	0.12
2 Clarification	262	6.17
3 Adequate answer	3923	92.35
4 Qualified answer	51	1.20
5 Inadequate answer	1	0.02
6 Don't know	6	0.14

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 December 2022 doi:10.20944/preprints202212.0369.v1

Table 7 – continued from previous page

Code	Respondent behavior counts	%
7 Refusal to answer	0	0.00

Various questionnaire item numbers and the number of respondents who answered with interruption (1) in all the three languages are shown in Table 8. The total number of persons who responded with an interruption for each item was not up to 15% or more. Question 1.B2 in section B, part 3 has a maximum of 2 respondents who answered with interruption.

Table 8. Proportion of respondents who answered with interruption (1) in all languages (English, Norwegian and Indonesia)

No.	Part	Section	Question No.	Count of Respondents #	Proportion (%)
1	Part 3	Section B	1.B2	2	6
2	Part 3	Section D	1a.B2	1	3
3	Part 5	Section A	1.A	1	3
3	Part 5	Section A	1.B2	1	3

Out of the 262 responses who sorted for clarifications, a total of 14 questionnaire items were answered by 15% or more respondents with clarification as shown in Table 9. Question number 1.A in part 4, Section A has the highest proportion of respondents (33%) who answered with clarification. Furthermore, questionnaire item number 1.B2 in part 4, Sect A, number 5.k in part 4, Sect B, and questionnaire item number 14 in part 4 section B each as the next highest proportion (22%)of respondents who answered with clarification.

Table 9. Proportion of respondents who answered with Clarification (2) in all languages (English, Norwegian and Indonesia)

No.	Part	Section	Question No.	Count of Respondents #	Proportion (%)
1	Part 1	Section A	4	6	17
2	Part 3	Section A	1.K	6	17
3	Part 4	Section A	1.K	6	17
4	Part 4	Section A	1.A	12	33
5	Part 4	Section A	1.B	7	19
6	Part 4	Section A	1.B2	8	22
7	Part 4	Section A	3.A	7	19
8	Part 4	Section B	4.K	6	17
9	Part 4	Section B	4.B	7	19
10	Part 4	Section B	5.K	8	22
11	Part 4	Section B	5.A	6	17
12	Part 4	Section B	5.B	7	19
13	Part 4	Section B	5.B2	6	17
14	Part 4	Section B	14	8	22

Table 10 consists of proportions of the respondents who did not answer exactly the respective questions in the administered questionnaires. Out of a total of 36 respondents who answered the questionnaires in English, Norwegian and Indonesian language, 15% or more of the respondents did

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 December 2022 doi:10.20944/preprints202212.0369.v1

not provide exact responses. A total of 22 questionnaire items were not adequately responded, out of a total of 118 questionnaire items. Questionnaire item 1.A in section A part 4 have the highest proposition (57%) of respondents who did not exactly respond to the item.

Table 10. More than 14% Proportion of respondents who did not answer exact in all languages (English, Norwegian and Indonesia)

No.	Part	**Section**	**Question No.**	**Count of Exact Respondents #**	**Count of Non-Exact Respondents #**	**Proportion (%)**
1	Part 1	Sect A	4	30	6	17
2	Part 3	Sect A	1.K	30	6	17
3	Part 3	Sect E	1	30	6	17
4	Part 4	Sect A	1.K	30	6	17
5	Part 4	Sect A	1.A	23	13	36
6	Part 4	Sect A	1.B	29	7	19
7	Part 4	Sect A	1.B2	27	9	25
8	Part 4	Sect A	3.A	29	7	19
9	Part 4	Sect B	4.K	26	10	28
10	Part 4	Sect B	4.B	29	7	19
11	Part 4	Sect B	5.K	27	9	25
12	Part 4	Sect B	5.A	29	7	19
13	Part 4	Sect B	5.B	29	7	19
14	Part 4	Sect B	5.B2	29	7	19
15	Part 5	Sect A	3a	29	7	19
16	Part 5	Sect B	14	28	8	22
17	Part 6	Sect A	10	29	7	19

Regarding response with qualification, out of the total number of 51 response behaviors which were counted as shown in Table 7 and Table 11 none of the questionnaire items got 15% or more of the 36 total number of respondents who answered with qualification. The high number of proportions of respondents who answered with qualification was about 11.1% and they responded to question 4.K in Part 4 and Section B and question 3a in Part 5 and section A.

Table 11. Proportion of respondents who answered with Qualification (4) in all languages (English, Norwegian and Indonesia)

No.	Part	**Section**	**Question No.**	**Count of Respondents #**	**Proportion (%)**
1	Part 1	Sect A	1	3	8.3
2	Part 1	Sect A	2	1	2.8
3	Part 1	Sect A	7	2	5.6
4	Part 3	Sect B	1.B	3	8.3
5	Part 3	Sect C	1.K	3	8.3
6	Part 3	Sect D	1	1	2.8
7	Part 3	Sect E	1	2	5.6
8	Part 4	Sect A	1.A	1	2.8
9	Part 4	Sect B	1.K	1	2.8
10	Part 4	Sect B	2.B	3	8.3

 doi:10.20944/preprints202212.0369.v1

Table 11 – continued from previous page

No.	Part	**Section**	**Question No.**	**Count of Respondents #**	**Proportion (%)**
11	Part 4	Sect B	3.A	1	2.8
12	Part 4	Sect B	4.K	4	11.1
13	Part 4	Sect B	5.K	1	2.8
14	Part 4	Sect B	5.A	1	2.8
15	Part 5	Sect A	1.A	1	2.8
16	Part 5	Sect A	3a	4	11.1
17	Part 5	Sect B	3	2	5.6
18	Part 5	Sect B	4	3	8.3
19	Part 5	Sect B	6	2	5.6
20	Part 5	Sect B	8	2	5.6
21	Part 5	Sect B	9	1	2.8
22	Part 5	Sect B	10	2	5.6
23	Part 5	Sect B	12	1	2.8
24	Part 5	Sect B	16	1	2.8
25	Part 5	Sect B	17	2	5.6
26	Part 6	Sect A	9	1	2.8
27	Part 6	Sect A	10	2	5.6

Table 12 also presents the proportion of respondents who answered with the "Don't know" code (6). From the responses, none of the questionnaire items also have 15% responses out of the total responses.

Table 12. Proportion of respondents who answered with Don't Know (6) in all languages (English, Norwegian and Indonesia)

No.	Part	**Section**	**Question No.**	**Count of Respondents #**	**Proportion (%)**
1	Part 4	Section A	1.B2	1	3
2	Part 4	Sect B	1.K	1	3
3	Part 4	Section B	5.B2	1	3
4	Part 5	Sect B	3	1	3
5	Part 5	Sect B	4	1	3
6	Part 5	Sect B	11	1	3

5.1.2. Presentation of results based on the languages

As the questionnaire was administered in 3 languages, a gist of the results is therefore presented in this section in those 3 languages thus English, Indonesian language and Norwegian language. Overall, 19 persons responded to the questionnaire in the Indonesian language, 10 persons responded in English, and 7 persons answered in Norwegian, bringing the total number of respondents to be 36.

With regards to responses in the Indonesian language, the 118 questionnaire items were exactly read to all the respondents. However not all the responses were exact. Questionnaire items that were not answered with exact by more than 14% of the respondents are shown in Table 13. From this table, 10 questionnaire items were not exactly respondents by 15% or more of the respondents. Questionnaire number 1.A in part 4, section A, 1.B in part 4 and section A and 5.k in Part 4 and Section B showed the highest proportion of respondents who did not answer exactly.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 December 2022 doi:10.20944/preprints202212.0369.v1

Table 13. More than 14% Proportion of respondents In Indonesia who did not answer exact

No.	Part	Section	Question No.	Count of Exact Respondents #	Count of Non-Exact Respondents #	Proportion (%)
1	Part 1	Section A	4	13	6	17
2	Part 4	Section A	1.K	13	6	17
3	Part 4	Section A	1.B2	13	6	17
4	Part 4	Section A	1.A	12	7	19
5	Part 4	Section A	1.B	12	7	19
6	Part 4	Section B	4.K	13	6	17
7	Part 4	Section B	5.K	12	7	19
8	Part 4	Section B	5.A	13	6	17
9	Part 4	Section B	5.B	13	6	17
10	Part 4	Section B	5.B2	13	6	17

With regards to English responses, Table 14 showed 31 questionnaire items in which 15% or more of the respondents did not respond with exact (3).

Table 14. More than 14% Proportion of respondents In English who did not answer exact

No	Part	Section	Question No.	Count of Exact Respondents #	Count of Non-Exact Respondents #	Proportion (%)
1	Part 1	Section A	1	7	3	30
2	Part 1	Section B	8	8	2	20
3	Part 2	Section A	9	8	2	20
4	Part 2	Section A	10	7	3	30
5	Part 3	Section A	1.K	7	3	30
6	Part 3	Section B	1.K	8	2	20
7	Part 3	Section B	1.B	8	2	20
8	Part 3	Section B	1.B2	6	4	40
9	Part 3	Section C	1.K	7	3	30
10	Part 3	Section C	2.B	7	3	30
11	Part 3	Section D	1	8	2	20
12	Part 3	Section E	1	6	4	40
13	Part 4	Section A	1.A	6	4	40
14	Part 4	Section A	1.B2	8	2	20
15	Part 4	Section A	2.K	7	3	30
16	Part 4	Section A	2.A	6	4	40
17	Part 4	Section A	1.B2	8	2	20
18	Part 4	Section A	3.B	8	2	20
19	Part 4	Section B	1.K	8	2	20
20	Part 4	Section B	4.K	8	2	20
21	Part 4	Section B	4.A	8	2	20
22	Part 5	Section A	3a	7	3	30
23	Part 5	Section B	3	8	2	20
24	Part 5	Section B	4	7	3	30
25	Part 5	Section B	6	7	3	30

 doi:10.20944/preprints202212.0369.v1

Table 14 – continued from previous page

No	Part	Section	Question No.	Count of Exact Respondents #	Count of Non-Exact Respondents #	Proportion (%)
26	Part 5	Section B	11	8	2	20
27	Part 5	Section B	14	7	3	30
28	Part 6	Section A	3	7	3	30
29	Part 6	Section A	6	8	2	20
30	Part 6	Section A	7	8	2	20
31	Part 6	Section A	10	7	3	30

With regards to the questionnaire item in Norwegian, 7 persons responded to the questionnaire, and out of the 118 questionnaire items, all of them were exactly read to the respondents. However, 11 questionnaire items were not exactly respondents by 2 or more respondents as shown in Table 15. This means that over 14% of the respondents of the Norwegian version of the questionnaire did not give straight-away answers to 11 of the questionnaire items.

Table 15. More than 14% Proportion of respondents (Norwegian) who did not answer exact

No.	Part	Section	Question No.	Count of Exact Respondents #	Count of Non-Exact Respondents #	Proportion (%)
1	Part 1	Section A	7	5	2	29
2	Part 1	Section A	8	5	2	29
3	Part 4	Section A	1.A	5	2	29
4	Part 4	Section B	2.B	4	3	43
5	Part 4	Section B	3.B2	5	2	29
6	Part 4	Section B	4.K	5	2	29
7	Part 5	Section B	1.A	5	2	29
8	Part 5	Section B	3	4	3	43
9	Part 5	Section B	14	5	2	29
10	Part 5	Section B	16	5	2	29
11	Part 6	Section A	10	5	2	29

5.1.3. Summary of read

With regards to the reading of the questionnaire which was responded in all the three languages (English, Norwegian, and Indonesian), out of 36 times that the questionnaire was read to the 36 respondents, 5 questionnaire items (out of 118) were not precisely read in more than 3 times (representing more than 10% of the total number of reads) as shown in Table 16. In fact, one questionnaire item (Question No. 1) in part 3, section E was not precisely read for more than 5 times representing more than 14% of the total reads.

Table 16. More than 10% of readers who did not read exact in all the languages (English, Norwegian and Indonesian)

No.	Part	Section	Question No.	Count of Exact Respondents #	Count of Non-Exact Respondents #	Proportion (%)

 doi:10.20944/preprints202212.0369.v1

Table 16 – continued from previous page

No.	Part	Section	Question No.	Count of Exact Respondents #	Count of Non-Exact Respondents #	Proportion (%)
1	Part 3	Section C	2.A	31	5	13.8
2	Part 3	Section C	2.B	32	4	11.1
3	Part 3	Section E	1	30	6	16.6
4	Part 5	Section A	1.A	32	4	11.1
5	Part 5	Section B	13	31	5	13.8

Out of the total of the 118 questionnaire items, 15 of them were not read exactly to 2 or more respondents, representing over 14% of the non-exact reads as shown in Table 17.

Table 17. More than 14% Proportion of readings in English of which the reading of the questionnaire items were not exact

No.	Part	Section	Question No.	Count of Exact Readings #	Count of Non-Exact Reading #	Proportion (%)
1	Part 1	Section A	2	7	3	30
2	Part 3	Section B	1.B2	7	3	30
3	Part 3	Section C	2.K	8	2	20
4	Part 3	Section C	2.A	5	5	50
5	Part 3	Section C	2.B	6	4	40
6	Part 3	Section C	2.B2	6	4	30
7	Part 3	Section D	1	8	2	20
8	Part 3	Section E	1	4	6	60
9	Part 3	Section E	1a.B	8	2	20
10	Part 4	Section A	1.K	8	2	20
11	Part 5	Section A	1.A	6	4	40
12	Part 5	Section A	1.B	8	2	20
13	Part 5	Section B	3a	8	2	20
14	Part 5	Section b	3	6	4	30
15	Part 5	Section B	13	6	4	40

6. Discussion of questionnaire design and pretesting methods

In strengthening security-conscious care behavior among healthcare workers, a study was initiated to assess the security practice level in various hospitals in Norway, Indonesia, and Ghana. However, healthcare workers in these countries consist of various user groups with varying information security responsibilities such as management level, end-users, and all user groups. The all user group primarily intersects both management and end-users. These groups of personnel have different knowledge gaps in terms of their understanding of information technology and information security with varying responsibilities. For instance, the management group of these healthcare staff has a deeper understanding of information technology and security with higher responsibilities to developing policies and ensuring their implementations in healthcare systems. On the contrarily, other users such as doctors and nurses may also just have the basic knowledge of healthcare information technology systems (such as EHR) to enable them to use these systems for therapeutic functions. In developing a questionnaire for analyzing the security practice of healthcare workers in a hospital which consists of all these user categories, there is hence the need to scale down the understanding of the questionnaire

 doi:10.20944/preprints202212.0369.v1

to meet all these user categories. Having pretested our questionnaire, this section discusses the findings in this study.

6.1. Principal findings

According to Hamed et al and Somekh et al., a quality questionnaire should be clear, unambiguous, and simple [79,80]. Additionally, questionnaire designers are to avoid asking more than one question in a single question and should not use negative and double negatives. Furthermore, a quality questionnaire should adopt the usage of mutually exclusive questions and avoid questions that may irritate respondents[79,80]. This can be achieved if the right methods are used in the design and pretesting stages of the questionnaire development. In this regard, questionnaire design and pretesting methods were explored and used in designing and pretesting for the questionnaire which is intended to be used for analyzing healthcare security practice. In brief, the identified methods for designing questionnaire includes rational method, prototypical, facet, construct, internal and external methods [15–18,21,22,24–29,64,65] as shown in 1 and Table 2. Additionally, the identified pretesting methods include conventional, cognitive interview, behaviour coding, debriefing, response latency, vignettes, experimental technique, focus group and expert reviews [33–36,38,39,41,44,46,49–53,71,77,78] method as shown in Table 4. These methods were assessed in this study, towards designing and pretesting questionnaire for healthcare security practice analysis modeling and incentivization (HSPAMI)[8–10] .

HSPAMI consists of multifaceted areas that are considered as constructs in modeling and analyzing healthcare staffs' security practice. This study includes security practices which are related to;

- Psychological influence on healthcare staff
- Social influence among healthcare staff
- Individual and organizational cultural effect
- Work demographic effect on healthcare staff security practice
- Social demographic influence on healthcare staff
- Healthcare staff security awareness and training

Additionally, there are various existing questionnaires for surveying for security practice in healthcare but which questionnaire design and pretesting methods are suitable for adopting existing items[72–74]? Moreover, the KAB (Knowledge, Attitude, and Behavior) approach [74,75] has been recommended to be used in the development of questionnaires in information security practice. This is because knowledge influences attitude and behavior and these need to be assessed with the information security practices. So this section analyses and discusses the gaps and objectives in each of the questionnaire design and pretesting methods and their suitability in adopting them in HSPAMI questionnaire design and pretesting.

6.2. Selection and adoption of the questionnaire design methods

The questionnaire design methods which were selected and adopted in this study are the rational, prototypical, facet, and internal methods as shown in Table 5.

Considering the rational method, a questionnaire that is designed with this method is incorporated with only the subjective view of the questionnaire developer. Therefore, a questionnaire that is assessed with the rational method alone can be suitable for exploration towards further studies but might not be effective enough for empirical studies for decision making such as incentivization measures, which is aimed towards improving security practice in healthcare. So the rational method was considered in the questionnaire design together with other methods in designing the questionnaire for HSPAMI. The rational method was used in separating questionnaire items into a section known as general security and privacy practice. From the perspective of the designers, this section was not related to any of the constructs as shown in Appendix 1.

The prototypical method is known to improve cognitive processing and stimulates a better understanding of the questionnaire items[15,18]. This is so since the organization of the questionnaire

 doi:10.20944/preprints202212.0369.v1

items is always categorized based on the items that are more prototypical to each other. The prototypical approach relies on informal (knowledge based on rules of thumb or tricks of the trade) or tacit knowledge (eg., knowledge based on experience or the job training) [15,17,18,64,65]. It is, therefore, suitable for the specification of implicit ideas and analyzing complex concepts[18]. The construction procedures are more elaborated in the prototypical method than the rational method and have therefore been considered as an effective method[18,64] particularly for existing questions. The prototypical method was therefore used to select relevant existing questionnaire items into various constructs of the HSPAMI study questionnaire. This was necessary since the existing questionnaire items were already tested and were needing modifications to fit into the various constructs of the study[75,76].

The facet method has been considered to be important in identifying comprehensive questionnaire items to cover all areas of the constructs but it only relies on formal knowledge of the construct [18]. This happened to suit well with the nomological framework of our study as shown in figure 1.

The facet method is related to the construct method in the sense that, both of them rely on formal knowledge [15,18] But the construct method requires theoretical and formal knowledge which can be used to develop a nomological network for empirical testing of hypothesis [18].

The method which seeks to address issues of internal validity is the internal method and it is suitable to be used for improving on an existing questionnaire. However, the specification and judgment of the items on the questionnaire are limited in scope because the method uses a statistical approach and not a theoretical framework that has the tendency to relate with a wider scope of variables. With regards to the external method, much as the wider scope of the study area can be covered, the idea of heterogeneous content has introduced complexity in using the method.

6.3. Selection and usage of the pretesting methods

6.4. Pretesting of questionnaire

Based on the guidance of the methods, a questionnaire for healthcare security practice was designed and pretested. The objective of the questionnaire was to collect the security behavior of healthcare workers by relating their characteristics with various information security practices. The healthcare workers that were targeted include healthcare professionals(doctors, nurses, laboratory and pharmacy personnel), healthcare administrators or secretaries, IT staff, and all end users who access the health information systems in their various healthcare organizations. As the targeted users consist of both highly skilled IT personnel and non-technical users, the questionnaire was intended to be designed and tested to suit all of these respondents. The characteristics of the healthcare workers which were being considered in this questionnaire include psychological traits, social and cultural characteristics. Staff demographic characteristics such as gender and work experience were also considered. Work-related demographics including roles, workload, and work emergency were as well considered in this study. So this study intends to relate all these staff characteristics with their security practices such as internet use, password management, social media use, information handling, and other security practices which are more prone to security violations of the healthcare workers.

A number of issues were identified when the questionnaire was pretested with conventional pretesting and behavior coding. With conventional pretesting, the respondents identified various problems with the questionnaire including issues relating to questions scale arrangement, incomplete response options, lack of understanding of the structure of information systems of respondents, use of undefined abbreviations, unrelated questions to respondents, unclear and complex questions, the insignificant difference between questions grammatical and spelling mistakes and security and privacy concerns of respondents.

For example, in a questionnaire item such as, "It is inconvenient to my work for me to use strong password policies such as multi-factor authentication, long password characters mixed with alphabets, numbers, and characters", it was realized that IT personnel in healthcare had no problem with understanding of the term multi-factor authentication. However, ordinary users of healthcare

 doi:10.20944/preprints202212.0369.v1

information systems including nurses, doctors, and healthcare administrators, who do not have technical skills in IT, could not understand the meaning of "multi-factor authentication". So in the repair of this question, an explanation of the term was provided. Additionally, as shown in Figure A1, a respondent indicated some of the questions which were unclear, and further indicated the exclusion of some of the roles in the laboratory department.

Related studies including Developing a Security Behavior Intentions Scale (SeBIS) and human aspect of information security questionnaire(HAIS-Q) were much focused on improving the effectiveness of surveys relating to information security. However, a number of shortfalls were observed. First of all, their target respondents did not consider comprehensive respondents of various roles and activities in an organizational setting like healthcare. Therefore, roles such as IT personnel who perform IT administrative duties including backups and updates were excluded. Additionally, though HAIS-Q used some pretesting methods, other pretesting methods were not assessed. Additionally, unlike this paper, the studies did not fully consider comprehensive questionnaire design methods.

6.5. Behavior coding responses

In the behavior coding exercise, 50 questionnaire items out of the 118 were remarked to have issues as shown in Table A1. Additionally, all the non paradigmatic responses in this study (interruption, clarification, qualification and don't know) except refusal to answer reported various questionnaire items to have issues as shown in Table 7, Table 8, Table 9, Table 10, Table 11, and Table 12. This implies that the questionnaire would have been administered with these problems despite having pretested it with the conventional method. In the conventional pretesting, the respondents answered these questionnaires based on their subjective understanding. Even if they had issues with any of the questionnaire items, the respondents in the conventional pretesting were not subjected to prop for clarifications prior to responding to each of the items and therefore answered the questions in their way of understanding. The behavior coding has effectively solved this gap by providing the respondents with the opportunity to inquire if they have any issues with any of the questionnaire items.

Much as the behavior coding indicated the issues pertaining to various questionnaire items, there have been suggestions that when 15% or more of the responses to the questionnaire items are non-paradigmatic, it is an indication that there are significant issues with that questionnaire item [41,49]. On that note, questionnaire items in Table 9 and Table 10 definitely needed significant repairs.

Furthermore, since the study was conducted in three languages, (English, Norwegian and Indonesian Language) translation errors in the questionnaire could not be overlooked. So the analysis of the responses in these languages also showed various errors in the questionnaire items. Table 13, Table14, and 15 respectively showed all the questionnaire items in Indonesian, English and Norwegian languages which significantly require repairs. The English language version showed more of its questionnaire items (31) that need repairs as compared to Indonesian(10) and Norwegian (11), suggesting that the English version comparatively had more issues with its questionnaire items.

As behavior coding, merely pinpoints to the questionnaire items that have issues without providing their causes [**?**], we augmented the behavior coding method by providing opportunities for respondents and interviewers to comment or indicate the cause of problems with each of the questionnaire items. A total of 50 questionnaire items out of the 118 attracted comments indicating various causes of the questionnaire item problems. This adjustment of the behavior coding strengthened the study and improved the quality of the questionnaire pretesting.

The behavior coding was adopted with the ultimate aim to pretest the questionnaire. So the interviewer's reading behavior was captured to be analyzed towards improving on the questionnaire. Some attempts have related the behavior codes to certain problems[41,49]. Questionnaire items that are not exactly read as worded have been phrased in an awkward way. Sometimes those items could also be having words that are difficult to pronounce. Also, questionnaire items that are frequently

 doi:10.20944/preprints202212.0369.v1

responded to with interruptions tend to provide infirm explanations at their conclusion. Additionally, responses to questionnaire items having sorted for clarification imply that the questionnaire item does not fit the respondents' experience or frame of reference and is often vague or contains a poorly defined term or concept. Whilst questionnaire items that are responded to with inadequate answers often require a level of details the respondents are not capable of providing.

7. Conclusions

While healthcare organizations are mandated to collect detailed sensitive personal health information for the purpose of correct patient identification and effective therapeutic measures, data breaches have become rampant with the potential to jeopardize effective healthcare. The adversaries usually gain access to the healthcare data through the human elements including the healthcare staff who have legitimate accesses to the healthcare records. To curtail this, research is being conducted using a questionnaire to determine the gaps between required security practice and current security practices in healthcare. The challenges of healthcare staff in complying with the security requirement would also be determined towards resolving any identified dilemmas. With such high ambition, an effective questionnaire needs to be developed to meet the needs of the diverse respondents who have variant IT skills gaps, experiences, and roles. Therefore, questionnaire design and pretesting methods were identified in a literature survey and used in designing and pretesting the questionnaire for healthcare security practice analysis.

The identified questionnaire pretesting methods were rational, prototypical, facet, construct, internal and external methods. Additionally, the pretesting methods which were identified include a conventional, cognitive interview, behavior coding, debriefing, response latency, vignettes, experimental, focus groups, and expert reviews.

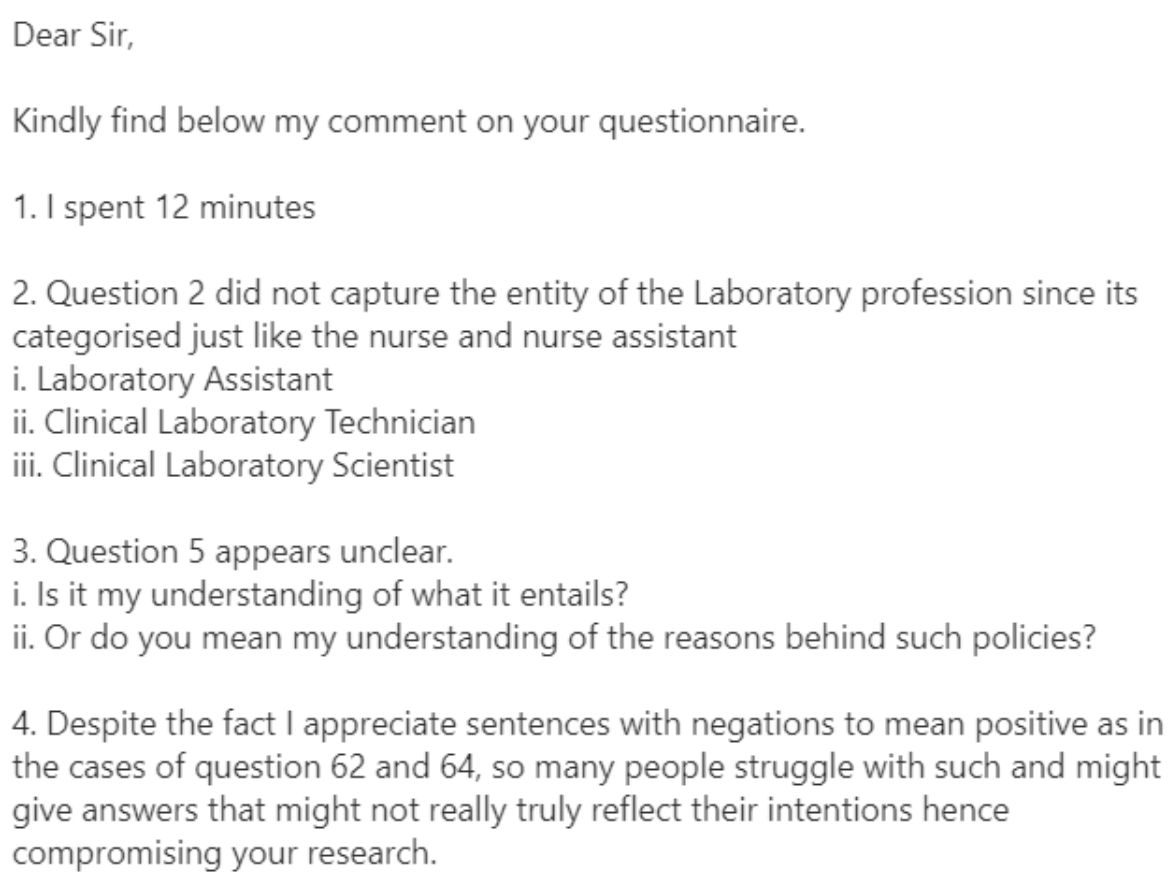
The contribution in this study is in different facets. Firstly, different types of questionnaire pretesting methods were identified in the survey. Secondly, the pretesting methods were accessed based on their advantages for effective pretesting of questionnaires towards healthcare security practice. Finally, effective pretesting methods were used to pretest questionnaires for healthcare security practice.

In the pretesting of the questionnaire, various questionnaire issues were identified to include issues with an arrangement of questions scales, lack of understanding of healthcare information systems' structure of all hospitals, unclear questions, the insignificant difference between questions, issues of security and privacy of respondents, problematic questions, inadequate questions, and complex terms. These issues were identified, having combined conventional pretesting method to cognitive interview and behavior coding. All the identified issues are pitfalls to watch when developing a questionnaire to be responded by an organization which is characterized with varying level of IT skills. Even though the identified problems have been corrected based on inputs from the study respondents and focus group experts, further works including animated and cartoonist questionnaire is to be explored in future studies. If a question in the questionnaire item mentions a term such as a "URL" or a "multi-factor authentication", an animated image or cartoon will be used to provide a better understanding of such wording.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 20 December 2022 doi:10.20944/preprints202212.0369.v1

Appendix A.

Appendix A.1. A sample responses from conventional pretesting


Dear Sir,

Kindly find below my comment on your questionnaire.

1. I spent 12 minutes

2. Question 2 did not capture the entity of the Laboratory profession since its categorised just like the nurse and nurse assistant
i. Laboratory Assistant
ii. Clinical Laboratory Technician
iii. Clinical Laboratory Scientist

3. Question 5 appears unclear.
i. Is it my understanding of what it entails?
ii. Or do you mean my understanding of the reasons behind such policies?

4. Despite the fact I appreciate sentences with negations to mean positive as in the cases of question 62 and 64, so many people struggle with such and might give answers that might not really truly reflect their intentions hence compromising your research.

Figure A1. Sample respond from a respondent in the pretesting study.

Table A1. Comments from behaviour coding

No	Part	Section	Question No.	Comments
1	Part 1	Sect A	1	What is your gender(add others)
2	Part 1	Sect A	2	Typo, personnel
3	Part 1	Sect A	7	Had formal training on i formation security ;Knows they have a policy but have not been trained;Respondent doesnt work in the hospital as a healthcare professional
4	Part 1	Sect B	8	Whant kind of brak?
5	Part 1	Sect B	10	specific on often
6	Part 2	Sect A	10	Complex question; Start with During emergency situation
7	Part 3	Sect A	1.K	is it about me in particula
8	Part 3	Sect B	1.K	can be harmful
9	Part 3	Sect B	1.B	update to social instead;THEIR NETWORK DOESNT ALLOW THAT

 doi:10.20944/preprints202212.0369.v1

Table A1 – continued from previous page

No	Part	Section	Question No	Comments
10	Part 3	Sect B	1.B2	Find out about security practice first before asking question. Put training aspect to it
11	Part 3	Sect C	1.K	Indicate Where;How to incoporate "i DONT KNOW". 3; Condition if user recieves alert messages or not options;Indicate where yes,no, cant remember or dont know
12	Part 3	Sect C	1.B2	typo
13	Part 3	Sect C	2.K	policy or etc; typo
14	Part 3	Sect C	2.A	Remove s from link
15	Part 3	Sect D	1	Restarts machine when instructed by the leader
16	Part 3	Sect E	1	incomplete caption, Typos; Work colleagues?
17	Part 3	Sect E	1a.K	Getting second opinion from colleague on results should be a problem however, sharing patient information to colleagues is bad
18	Part 3	Sect E	1a.B	From seeing instead
19	Part 4	Sect A	1.A	choices are confusing
20	Part 4	Sect A	1.B2	not applicable to her
21	Part 4	Sect A	2.K	Add patients information. Question is too general
22	Part 4	Sect A	2.B	what is anything
23	Part 4	Sect A	2.B2	change dont to did not, OR didn't
24	Part 4	Sect A	3.K	punished by who?
25	Part 4	Sect A	3.A	Type of action taken(we need to specify if I leave somebody will access my data)
26	Part 4	Sect A	3.B	Giving your password to a friend to view patient data : can be considered
27	Part 4	Sect B	1.A	Choices are confusing
28	Part 4	Sect B	1.B2	Not applicable
29	Part 4	Sect B	2.K	Clarrification;Require "not applicable"
30	Part 4	Sect B	2.B	Not application
31	Part 4	Sect B	3.B2	Not applicable
32	Part 4	Sect B	4.B	What is a shared password?
33	Part 4	Sect B	5.A	Not applicable
34	Part 4	Sect B	5.B	how do you dispose the sensitive
35	Part 4	Sect B	5.B2	Many typos;Not applicable
36	Part 5	Sect A	1.A	Typos (Spelling error: Thrust, sase)
37	Part 5	Sect A	3a	use check box;Should ask this earlier;.had formal training
38	Part 5	Sect B	3	Remove if
39	Part 5	Sect B	4	Introduce Which,; management team; Not all the time

 doi:10.20944/preprints202212.0369.v1

Table A1 – continued from previous page

No	Part	Section	Question No	Comments
40	Part 5	Sect B	6	use "someone"
41	Part 5	Sect B	8	personal mobile, add some example
42	Part 5	Sect B	9	Qualified answer
43	Part 5	Sect B	10	Temporal -check;what is sikkerhetstiltak'
44	Part 5	Sect B	13	Remove if
45	Part 5	Sect B	14	which media
46	Part 5	Sect B	16	Not clear question
47	Part 6	Sect A	3	clarify dependable
48	Part 6	Sect A	6	contradicts extrovert
49	Part 6	Sect A	9	Clarify calm
50	Part 6	Sect A	10	Clarify conventional;Use a different word for 'konvensjonal'

References

1. Pedersen, S. and G. Hartvigsen, Lessons learned from 25 years with telemedicine in Northern Norway. 2015.
2. Drye, S. and A. Salam, A framework for health care information assurance policy and compliance. Commun. ACM, 2010. 53: p. 126-131
3. Williams, P.A., 2008. In a 'trusting'environment, everyone is responsible for information security. Information Security Technical Report, 13(4), pp.207-215.
4. Box, D. and Pottas, D., 2013. Improving information security behaviour in the healthcare context. Procedia Technology, 9(2013), pp.1093-1103.
5. SearchHealthIT. Hospital takes aim at patient health data security with AI tools. 2019; Available from: https://searchhealthit.techtarget.com/feature/Hospital-takes-aim-at-patient-health-data-security-with-AI-tools.
6. HealthITSecurity, Incentivize Cybersecurity Best Practices for Data Security. 2017, @SecurityHIT.
7. Ehrenfeld, J.M., Wannacry, cybersecurity and health information technology: A time to act. Journal of medical systems, 2017. 41(7): p. 104.
8. Yeng, P.K., B. Yang, and E.A. Snekkenes. Framework for Healthcare Security Practice Analysis, Modeling and Incentivization. in International Workshop on Big Data Analytics for Cyber Threat Hunting. 2019. Los Angeles: IEEE Big Data.
9. Yeng, P., B. Yang, and E. Snekkenes. Observational Measures for Effective Profiling of Healthcare Staffs' Security Practices. in 2019 IEEE 43rd Annual Computer Software and Applications Conference (COMPSAC). 2019.
10. Kandabongee Yeng, P., B. Yang, and E. Arthur Snekkenes, Healthcare Staffs' Information Security Practices Towards Mitigating Data Breaches: A Literature Survey. Studies in health technology and informatics, 2019. 261: p. 239-245.
11. Yeng, P., et al., Data-Driven and Artificial Intelligence (AI) Approach for Modelling and Analyzing Healthcare Security Practice: A Systematic Review. 2020.
12. e-helse, D.f., Code of conduct for information security and data protection in the healthcare and care services sector, D.f. e-helse, Editor. 2018.
13. Verison. Data Breaches Report. 2019. Accessed: https://www.nist.gov/system/files/documents/2019/10/16/1-2-dbir-widup.p
14. Drye, S. and A. Salam, A framework for health care information assurance policy and compliance. Commun. ACM, 2010. 53: p. 126-131
15. Oosterveld P, Vorst HC, Smits N. Methods for questionnaire design: a taxonomy linking procedures to test goals. Quality of Life Research. 2019 Sep 15;28(9):2501-12.
16. Simms LJ. Classical and modern methods of psychological scale construction. Social and Personality Psychology Compass. 2008 Jan;2(1):414-33.

17. Waller NG, DeYoung CG, Bouchard Jr TJ. The recaptured scale technique: A method for testing the structural robustness of personality scales. Multivariate behavioral research. 2016 Jul 3;51(4):433-45.
18. OOSTERVELD, Paul, et al. Questionnaire design methods. Berkhout Nijmegen BV, 1996.
19. Clark EV. Cognitive development and the acquisition of language. What's in a word? On the child's acquisition of semantics in his first language. 1973.
20. Rosch, E. H. (1978). Principles of categorization. In E. H. Rosch & D. B. Lloyd (Eds.), Cognition and categorization (pp. 27–48). Hillsdale, NJ: Erlbaum.
21. Guttman L. An outline of some new methodology for social research. Public Opinion Quarterly. 1954 Jan 1;18(4):395-404.
22. Guttman L. Introduction to facet design and analysis', in proceeding the Fifteenth international congress of psychology.
23. Stouthard ME, Mellenbergh GJ, Hoogstraten J. Assessment of dental anxiety: a facet approach. Anxiety, Stress and Coping. 1993 Jan 1;6(2):89-105.
24. Jackson, D. N. Structural personality assessment. 1973. In B. B. Wolman (Ed.), Handbook of general psychology (pp. 775–792). Englewood Cliffs, NJ: Prentice Hall.
25. Jackson, D. N. The dynamics of structured personality tests.1971. Psychological Review, 78(3), 229.
26. Cattell, R. B. , Saunders, D. R., & Stice, G. Handbook for the sixteen personality factor questionnaire. 1957. '16 P. F. test' forms A, B, and C. Champaign, IL: IPAT.
27. Comrey, A. L. . Factor-analytic methods of scale development in personality and clinical psychology. Journal of Consulting and Clinical Psychology. 1988. 56, 754–761.
28. Thurstone, L. L. (1931). Multiple factor analysis. Psychological Review, 38, 406–427.
29. Meehl, P. E. The dynamics of "structured" personality tests. Journal of Clinical Psychology, 1(4), 296–303. Reprinted in D. N. Jackson and S. Messick (Eds.), (1967). Problems in human assessment (pp. 517–522).
30. de Jong PF. An application of the prototype scale construction strategy to the assessment of student motivation. Journal of personality. 1988 Sep;56(3):487-508.
31. Amelang M, Herboth G, Oefner I. A prototype strategy for the construction of a creativity scale. European Journal of Personality. 1991 Sep;5(4):261-85.
32. Edwards, Allen L. "The measurement of personality traits by scales and inventories." (1970).
33. Presser S, Couper MP, Lessler JT, Martin E, Martin J, Rothgeb JM, Singer E. Methods for testing and evaluating survey questions. Public opinion quarterly. 2004 Mar 1;68(1):109-30.
34. Cannell CF, Kahn RL. The collection of data by interviewing. Research methods in the behavioral sciences. 1953:327-80.
35. Moser CA, Kalton G. Survey methods in social investigation. Routledge; 2017 Mar 2.
36. Converse JM, Jean McDonnell C, Presser S. Survey questions: Handcrafting the standardized questionnaire. Sage; 1986.
37. Martin E, Schechter S, Tucker C. Interagency collaboration among the cognitive laboratories: past efforts and future opportunities. Statistical Policy Working Paper 28: 1998 Seminar on Interagency Coordination and Cooperation.
38. Schechter S, Beatty P, Block A. Cognitive issues and methodological implications in the development and testing of a traffic safety questionnaire. InProceedings of the Survey Research Methods Section of the American Statistical Association 1994 (pp. 1215-1219).
39. Drennan, J. (2003). Cognitive interviewing: verbal data in the design and pretesting of questionnaires. Journal of advanced nursing, 42(1), 57-63.
40. Goldenberg KL. Using cognitive testing in the design of a business survey questionnaire. Inpresentation at the annual meeting of the American Association for Public Opinion Research, Salt Lake City, UT 1996 May.
41. Czaja R. Questionnaire pretesting comes of age. Marketing Bulletin-Department of Marketing Massey University. 1998 May;9:52-66.
42. Taylor C. Clinical problem-solving in nursing: insights from the literature. Journal of Advanced Nursing. 2000 Apr;31(4):842-9.
43. Schuwirth LW, Verheggen MM, Van der Vleuten CP, Boshuizen HP, Dinant GJ. Do short cases elicit different thinking processes than factual knowledge questions do?. Medical Education. 2001 Apr 22;35(4):348-56
44. Willis, G., T. DeMaio, and B. Harris-Kojetin. "Cognition and survey research." (1999): 133-53.

45. Conrad F., Blair J. & Tracy E. (1999) Verbal reports are data! A theoretical approach to cognitive interviews. In Proceedings of the Federal Committee on Statistical Methodology Research Conference, Tuesday B Sessions, Arlington, VA, pp. 11– 20.
46. Forsyth B, Rothgeb JM, Willis GB. Does pretesting make a difference? An experimental test. Methods for testing and evaluating survey questionnaires. 2004 Jun 25:525-46.
47. Royston P, Bercini D. Questionnaire design research in a laboratory setting: results of testing cancer risk factor questions. In1987 Proceedings of the Section on Survey Research Methods 1987 (pp. 829-33).
48. Belson WA. The design and understanding of survey questions. Lexington Books; 1981.
49. Oksenberg L, Kalton G. New strategies for pretesting survey questions. Journal of official statistics. 1991 Sep 1;7(3):349–365
50. Van der Zouwen J, Smit JH. Evaluating survey questions by analyzing patterns of behavior codes and question-answer sequences: A diagnostic approach. Methods for testing and evaluating survey questionnaires. 2004 Jun 25:109-30.
51. Draisma S, Dijkstra W. Response latency and (para) linguistic expressions as indicators of response error. Methods for testing and evaluating survey questionnaires. 2004 Jun 25;131147.
52. Hughes KA. Comparing pretesting methods: Cognitive interviews, respondent debriefing, and behavior coding. Survey Methodology. 2004;2:1-20.
53. Biemer PP. Modeling measurement error to identify flawed questions. Methods for testing and evaluating survey questionnaires. 2004 Jun 25:225-46.
54. Reeve BB, Mâsse LC. Item response theory modeling for questionnaire evaluation. Methods for testing and evaluating survey questionnaires. 2004 Jun 25;1:247-74.
55. Beatty, Paul C., and Gordon B. Willis. "Research synthesis: The practice of cognitive interviewing." Public opinion quarterly 71.2 (2007): 287-311.
56. Bolton, Ruth N and Bronkhorst, Tina M.uestionnaire pretesting: Computer-assisted coding of concurrent protocols.Jossey-Bass:1996
57. Conrad, Frederick, Johnny Blair, and Elena Tracy. “Verbal Reports Are Data! A Theoretical Approach to Cognitive Interviews.” In Proceedings of the 1999 Federal Committee on Statistical Methodology Research Conference. Washington, DC: Office of Management and Budget.2000.
58. Beatty, Paul, Gordon Willis, and Susan Schechter. Evaluating the Generalizability of Cognitive Interview Findings. In Statistical Policy Working Paper No. 26: Seminar on Statistical Methodology in the Public Service. Washington, DC: Statistical Policy Office, U.S. Office of Management and Budget.1997.
59. Tucker C. Methodological issues surrounding the application of cognitive psychology in survey research. Bulletin of Sociological Methodology/Bulletin de Méthodologie Sociologique. 1997 Jun;55(1):67-92.
60. Ericsson KA, Simon HA. Verbal reports as data. Psychological review. 1980 May;87(3):215.
61. Lessler, Judith T., and Barbara H. Forsyth. "A coding system for appraising questionnaires." (1996).
62. Evans SC, Roberts MC, Keeley JW, Blossom JB, Amaro CM, Garcia AM, Stough CO, Canter KS, Robles R, Reed GM. Vignette methodologies for studying clinicians' decision-making: Validity, utility, and application in ICD-11 field studies. International journal of clinical and health psychology. 2015 May 1;15(2):160-70.
63. Royston P, Bercini D. Questionnaire design research in a laboratory setting: results of testing cancer risk factor questions. In1987 Proceedings of the Section on Survey Research Methods 1987 (pp. 829-33).
64. Broughton R. A prototype strategy for construction of personality scales. Journal of Personality and Social Psychology. 1984 Dec;47(6):1334.
65. Rosch EH. On the internal structure of perceptual and semantic categories. In Cognitive development and acquisition of language 1973 Jan 1 (pp. 111-144). Academic Press.
66. Radhakrishna RB. Tips for developing and testing questionnaires/instruments. Journal of extension. 2007 Feb;45(1):1-4.
67. Kazi AM, Khalid W. Questionnaire designing and validation. Journal of the Pakistan Medical Association. 2012;62(5):514.
68. Brown JD. What is construct validity. Available at jalt. org/test/bro_8. htm (accessed 10 June 2013). 2000.
69. Kaplan RM, Bush JW, Berry CC. Health status: types of validity and the index of well-being. Health services research. 1976;11(4):478.
70. Edelen MO, Reeve BB. Applying item response theory (IRT) modeling to questionnaire development, evaluation, and refinement. Quality of life research. 2007 Aug 1;16(1):5.

71. Krueger RA. Focus groups: A practical guide for applied research. Sage publications; 2014 Aug 14.
72. Fernandez-Aleman, J. L., Sanchez-Henarejos, A., Toval, A., Sanchez-Garcia, A. B., Hernandez-Hernandez, I., and Fernandez-Luque, L. (2015). Analysis of health professional security behaviors in a real clinical setting: an empirical study. Int J Med Inform, 84(6), 454-467. doi:10.1016/j.ijmedinf.2015.01.010
73. Ng BY, Kankanhalli A, Xu YC. Studying users' computer security behavior: A health belief perspective. Decision Support Systems. 2009 Mar 1;46(4):815-25.
74. Anwar, Mohd, et al. "Gender difference and employees' cybersecurity behaviors." Computers in Human Behavior 69 (2017): 437-443.
75. Parsons, K., et al., The Human Aspects of Information Security Questionnaire (HAIS-Q): Two further validation studies. Computers & Security, 2017. 66: p. 40-51.
76. Caldwell, A. HAIS-Q: A smart solution to cyber security. 2017 2017-08-30; Available from: https://www.dst.defence.gov.au/podcast/hais-q-smart-solution-cyber-security.
77. Yan T, Kreuter F, Tourangeau R. Evaluating survey questions: A comparison of methods. Journal of Official Statistics. 2012 Dec 1;28(4):503-29.
78. Forsyth, B.H. and Lessler, J.L. (1991). Cognitive Laboratory Methods: A Taxonomy. In Measurement Errors in Surveys, P. Biemer, R. Groves, L. Lyberg, N. Mathiowetz,and S. Sudman (eds). New York: John Wiley, 393–418.
79. Taherdoost, H., 2016. How to Design and Create an Effective Survey/Questionnaire; A Step by Step Guide. International Journal of Academic Research in Management (IJARM), 5(4), pp.37-41.
80. Somekh, B. and Lewin, C. eds., 2011. Theory and methods in social research. Sage.
81. GOSLING, Samuel D.; RENTFROW, Peter J.; SWANN JR, William B. A very brief measure of the Big-Five personality domains. Journal of Research in personality, 2003, 37.6: 504-528.
82. Fowler, F. J., Jr., & Cannell, C. F. (1996). Using behavioral coding to identify cognitive problems with survey questions. In N. Schwarz and S. Sudman (Eds.), Answering questions: Methodology for determining cognitive and communicative processes in survey research (p. 15–36). Jossey-Bass/Wiley.
83. Wil Dijkstra. (2011). Behavior Coding. In Using behavioral coding to identify cognitive problems with survey questionsh (p. 54–55). Sage Publications, Inc, DOI:https://dx.doi.org/10.4135/9781412963947.

Sample Availability: Samples of the compounds ..... are available from the authors.