

Business and Impact Alignment of Questionnaire

Kateryna Piterenko



Master's Thesis
Master of Science in Media Technology
30 ECTS
Department of Computer Science and Media Technology
Gjøvik University College,

Avdeling for
informatikk og medieteknikk
Høgskolen i Gjøvik
Postboks 191
2802 Gjøvik

Department of Computer Science
and Media Technology
Gjøvik University College
Box 191
N-2802 Gjøvik
Norway

Business and Impact Alignment of Questionnaire

Kateryna Piterenko

2013/06/02

Abstract

Survey is a quick common instrument for information gathering. One of the most common forms of the survey is a questionnaire. Questionnaire is “*a technique for gathering statistical information about the attributes, attitudes or actions of a population by administering standardized questions to some or all its members.*” [5] In this thesis we are discussing business impact of questionnaire and problems related to questionnaire design. Aim of this project is universal solution for questionnaire design. In our opinion, questionnaire design has appeared as the weakest link in the survey. Moreover, not everyone realizes that it causes frequently confusing results. We decided to design our own solution for that.

We propose a new methodology of questionnaire design. Our approach is based on start-point framework and testing/verification part of the questionnaire on its developing stage. Notion consists of possibility to test and verify questionnaire for design issues before carry it out to the target group. This option additionally helps to design question types, avoid different sources of bias, and sometimes correct and review the research questions. The proposed approach for testing and verification includes three closed loops of iterations, researcher needs to run to improve the questionnaire. One of the advantages of our approach, which we would like to highlight, is simplicity of iterations, which can be repeated constantly until the quality of the questionnaire will be satisfactory. Loops give researcher different options of how to improve the questionnaire. Loops can be used separately or all together. Then, depending on time resources, researcher can choose one or more options and, additionally, the number of the iterations. We believe that our methodology will enhance the questionnaire performance significantly, hence, the survey results quality in general.

To test our own methodology, the conducted survey at Gjøvik University College has been chosen. The survey was addressed to measure the effectiveness of information security awareness program [43]. The survey was held in 2011 within the employees of Gjøvik University College. The survey is a part of another master thesis, which has been done in collaboration with Norsk Senter for Informasjonssikring (NorSIS). In this master thesis, NorSIS has requested to perform a new evaluation of this survey to assume better results. In terms of this project, we provide a full re-evaluation of the survey: starting from questionnaire design and ending with results interpretation. For the re-evaluation, we use our own methodology to present it in working conditions. Evaluation of questionnaire design has an impact on the statistical analysis and influences the results. The statistical analysis is performed using IBM SPSS Statistics software solutions and addressed to measure effectiveness of the training programs and their separate performances. The re-evaluation results are presented with counted in impact of questionnaire design analysis.

Acknowledgments

I would like to thank the supervisor of this master thesis, Dr. Prof. Bernhard M. Hämmerli, for his complete support, inspiration and ideas. This work would not be possible without his guidance and supervision. He also helped me to start believe in myself and that everything is possible.

I also want to thank my dear Farhad for believing in me and never letting me to give up. Without him, being always next to me, nothing would be as it is.

Moreover, I want to thank my mom and sister for their support and faith.

Kateryna Pitereno, 2nd of June 2013

Contents

Abstract	iii
Acknowledgments	v
Contents	vii
List of Figures	ix
List of Tables	xi
1 Introduction	1
1.1 Topic Covered by the Project	1
1.2 Keywords	2
1.3 Problem Description	2
1.4 Justification, Motivation and Benefits	2
1.5 Research Questions	3
1.6 Claimed Contributions	3
1.7 Structure of Thesis	3
2 Choice of Methods	5
3 State of the Art	7
3.1 Background	7
3.2 Questionnaire: What are the Advantages?	10
4 Methodology	13
4.1 Methodology of Questionnaire Design	13
4.1.1 Framework	14
4.1.2 Testing and Verification	17
4.1.3 Cover Letter	20
5 Survey Evaluation: Effectiveness of Information Security Awareness Program	21
5.1 Qualitative Analysis of the Dataset	21
5.2 Questionnaire Design Evaluation	23
5.3 Statistical Analysis and Results Interpretation	35
6 Future Work	53
7 Conclusion	55
Bibliography	57
A Improvement of Questionnaire Design Using Semi-Structured Interviews	61
B Frequencies Distributions among Both Surveys	71
C Mean Plots	87

List of Figures

1	General scheme.	5
2	General components of information security.	10
3	Architecture of the Information Security Culture Decision Support System.	12
4	Complete scheme of the methodology.	15
5	Trusted group sampling.	19
6	Responses, first survey.	36
7	Responses, second survey.	36
8	Frequencies - group #1, first survey.	40
9	Frequencies - group #1, second survey.	40
10	Frequencies - group #2, first survey.	41
11	Frequencies - group #2, second survey.	41
12	Second survey: percentage between groups of respondents.	43
13	Descriptives - questions group #1.	43
14	Test of Homogeneity of Variances - questions group #1.	44
15	ANOVA table - questions group #1.	44
16	LSD - questions group #1.	45
17	Descriptives - questions group #2.	47
18	Test of Homogeneity of Variances - questions group #2.	48
19	ANOVA table - questions group #2.	49
20	LSD - questions group #2.	50
21	Contrast coefficients for participants and non-participants.	51
22	Contrast test for training participants and non-participants - questions group #1.	51
23	Contrast test for training participants and non-participants - questions group #2.	52
24	First group - Password Management.	77
25	Second group - Policy Regulations Awareness.	85
26	First group of questions (survey #2) - mean plots.	89
27	Second group of questions (survey #2) - mean plots.	93

List of Tables

1	Comparison of the questionnaires.	26
2	Questionnaire after first loop - semi-structured interviews.	29
3	Recoded questions from both questionnaires.	38
4	Grouping.	39
5	Questions for the first semi-structured interview.	62
6	Questions for the second semi-structured interview.	64
7	Questions for the third semi-structured interview.	65
8	Questions for the fourth semi-structured interview.	66
9	Questions for the fifth semi-structured interview.	67
10	Questions for the sixth semi-structured interview.	68
11	Questions for the seventh semi-structured interview.	70

1 Introduction

This chapter gives an introduction to the topic covered by this thesis. Here are presented topic of the project, problem description and stated research questions. We also claim the contributions in this chapter. By the end, we describe the structure of the report.

1.1 Topic Covered by the Project

The topic of this master thesis is evaluation of the business and alignment impact of the questionnaire. The ubiquity of the questionnaire survey attests to their utility. It has a geographic flexibility; it can be mailed out, so it does not depend on respondent's availability. Responders can fill it in whenever they have time. Moreover, questionnaire survey requires the lowest cost comparing with the other options to carry the survey out to the target group. Also, it has an anonymity of the respondent. Cover letter usually states that respondents' answers are confidential. Anonymity reduces social desirability bias. Hence, it makes questionnaire a really attractive way to conduct the survey.

However, nowadays researches often face a problem that survey results of the evaluation do not meet researchers' expectations and the quality of the results is unsatisfactory. Not everyone realizes that questionnaire design is a prolonged and arduous exercise. So, we asked ourselves, why questionnaire does not provide desirable outcome? The problem lies inside the questionnaire. Researchers often do not spend appropriate time on the questionnaire design, they concentrate more on the evaluation part and statistical analysis of the results. But it is obvious, that we cannot get good results as an output while providing a bad input - weak questionnaire design. Apparently, to assume good quality of the questionnaire, a methodology is needed.

Today, the survey and the questions are available. However, the evaluation is not satisfying and, in general, a new methodology for designing questionnaires should be given. This study is in collaboration with NorSIS. In this project, we create a new methodology of questionnaire design and give its own evaluation. To create a new methodology, many of the existing recommendations are investigated, analyzed, improved and modified. Based on our own experience, methodology takes a shape. To provide relevant verification of this methodology, existed survey at Gjøvik University College has been chosen. The survey was part of master thesis of Ilirjana Veseli in 2011 at Gjøvik University College in collaboration with NorSIS [43]. Questions of the survey were designed to measure employees' knowledge, understanding and compliance of the Gjøvik University College information security policy and related legal regulations [13].

Main issues, which led this survey to be chosen for the re-evaluation, are following: unsatisfactory analysis of the survey, in doubt questionnaire design, relevance of the data. Our methodology is addressed to verify questionnaire design and based on the verification, we can

give a part-answer related to the relevance of the following statistical analysis.

1.2 Keywords

Analysis, awareness, design, evaluation, information security, information security culture, methodology, statistics, SPSS, questionnaire, survey.

1.3 Problem Description

While using questionnaire survey in the research, researchers concentrate mostly on the evaluation stage. Exist incredible amount of literature, studies, methodologies, approaches, techniques and software solutions to perform statistical analysis of a data collected using a questionnaire. However, even while following all these rules and recommendations, extremely often results do not meet researches expectations. Overall, analysis shows results which cannot be taken into consideration as relevant ones. Such results, what is even worse, might cause confusing results. To check the relevance of obtained results, sample size is not enough. More likely, when facing such problem, everything related to the evaluation will be correct and clear, nevertheless, the results will not be trusted. Why is that? Here we can say because of the insufficient questionnaire design. It is clear that while providing wrong input, even through a perfect algorithm, it is impossible to obtain relevant output. Same with questionnaire and evaluation. Chosen techniques for evaluation are perfectly applied and used, but because of insufficient questionnaire, results fail.

Our project proposes a solution - a questionnaire design methodology. Solution must lead researcher while developing a questionnaire. It must give the hints and general rules of how to create a first-version questionnaire and, what is more important, how to test and verify questionnaire before carry it out. Methodology should not depend on concrete topic of the survey, it should be in-depth as possible

The main challenge is to create a methodology that makes possible, using easy tools and techniques, to improve the questionnaire to the efficient state, so the following evaluation process of the current survey can provide relevant results.

1.4 Justification, Motivation and Benefits

Questionnaire is a relatively cheap and suitable for the following statistical analysis way of collecting the data. Moreover, it is easy to carry it out to the responders. These factors make questionnaire common and widespread technique to perform a survey. Because questionnaire is so commonly used and preprocessing data gathering tool, improvement of its performance will not be left unnoticeable. It is used mostly in each research to gather data, to check some statements, to verify fundings and so on.

Let us make an example, each year at Gjøvik University College with just master theses projects performed around 10 questionnaire surveys. They all are addressed to measure and evaluate different cases; some of them are the aim of the projects, some of them are additional

research. This example is based on the projects within Faculty of Computer Science and Media Technology. And, unfortunately, most of these surveys were not so successful. Why is that? The answer is given above. For the same reason, we fail because we get wrong results to testify hypotheses, answer research questions, verify fundings etc. The hypotheses are formulated correctly, the research is designed efficient, the tools are chosen correctly, the evaluation is performed appropriate and results are still lead us to the wrong or, even worse, confusing, conclusions.

It is clear now that a methodology is needed. It will be extremely useful for researches, which are going on at Gjøvik University College, as well as out of the college, on higher level.

There is amount of different examples of questionnaire usage with information security. They are provided in the state-of-the-art chapter (3).

1.5 Research Questions

In order to create a new methodology of questionnaire design, we need to answer following questions:

- What are the existing approaches and recommendations for the questionnaire design?
- What should be done on the first stage of designing the questionnaire?
- How can we test and verify the questionnaire on the development stage?
- How should be questionnaire supported?
- When quality of the questionnaire is adequate?

1.6 Claimed Contributions

While investigating existing approaches and recommendations, which are addressed to design a questionnaire, we analyze and improve them. In addition, we introduce new way to maintain questionnaire design - semi-structured interviews, which are built on a questionnaire. All existing approaches claim their goal as questionnaire design only, no tests or verifications. We focus on verification and test of the questionnaire on the development stage. Because, even using promising set of rules to create a questionnaire, it should be tested and verified against biases, leading questions and a lot of other issues, which make a questionnaire a bad input for the survey.

Therefore, our methodology of questionnaire design contains three main frames: framework, testing methodology, and a subset of testing methodology - trusted group selection.

1.7 Structure of Thesis

The report of this thesis organized as follows: chapter 2 describes the methods which are used to answer research questions; chapter 3 describes state-of-the-art related to the research questions and particular importance and benefits of questionnaire; in chapter 4 a new methodology of

questionnaire design is given; chapter 5 contains re-evaluation of the survey using our methodology and statistical analysis; in chapter 6 future work is discussed and in chapter 7 conclusion is given.

2 Choice of Methods

In this thesis both qualitative and quantitative research methodologies are used. We apply qualitative approach to investigate existing recommendations for the questionnaire design, analyze and structure them out. This is a background for a new methodology of questionnaire design. For the evaluation part, we apply mixed quantitative and qualitative approaches. First, we analyze chosen questionnaire design regarding to our methodology, then, we perform possible improvements of questionnaire design, using testing and verification. Afterwards, we use statistical analysis to re-evaluate survey [34].

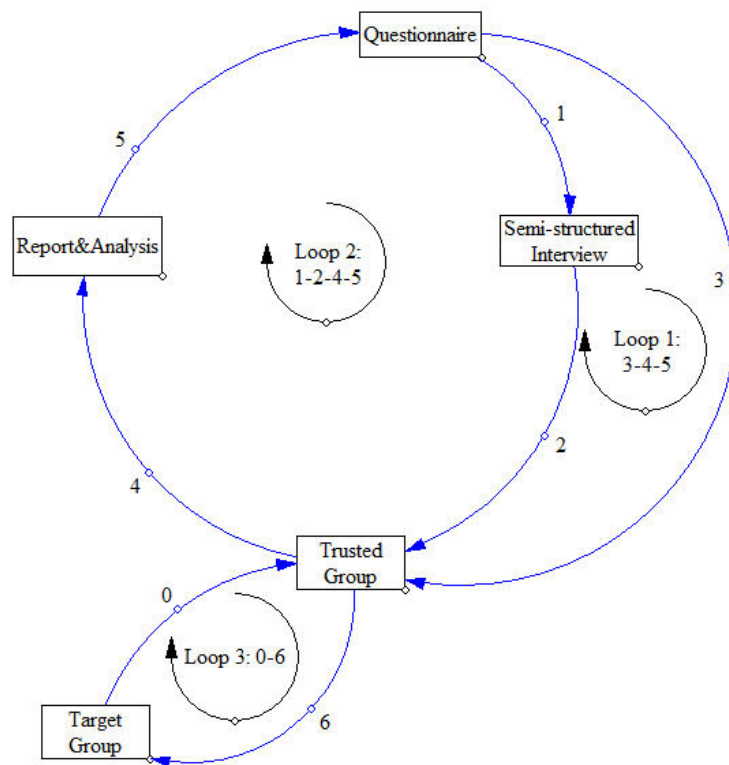


Figure 1: General scheme.

Strategy of the new methodology of questionnaire design bases on the research questions. The methodology should answer research questions. It must cover, not exclude, already existing recommendations and hints for questionnaire design, supporting tools, e. g. cover letter; it must give the possibility to test and verify questionnaire on its development stage; it should be able

to identify when questionnaire's quality is sufficient. Additionally, it should include tools and/or guidelines for testing and verification. For the following work, we designed general scheme that describes processes in the methodology (fig.1). The processes are next:

loop 1 is testing questionnaire through the semi-structured interviews, which are built on the questions from the questionnaire, through the trusted group,

loop 2 is verifying the questionnaire using trusted groups,

loop 3 selects trusted groups.

To complete statements given above, we chose next research strategies:

- literature overview of existing approaches for the questionnaire design,
- developing logical components of the methodology of questionnaire design,
- testing own methodology on already designed survey, this testing is addressed to accomplish two goals: test the methodology and estimate questionnaire design of the survey chosen for the re-evaluation,
- statistical analysis to complete re-evaluation of chosen survey.

Briefly speaking, we designed our research strategy to obtain maximum performance from each subtask. At the same time, we divide our research into subtasks and keep them interconnected. Like re-evaluation, for example. The aim of it is to assume better results of information security awareness survey. So, before statistical analysis, we use qualitative analysis of questionnaire design. What we get from this qualitative analysis, besides analysis results, is presentation of our methodology in working conditions. This part of evaluation includes semi-structured interviews with Gjøvik University College, questionnaire improvement, carrying questionnaire out to the trusted group.

Applying current strategy, we seek to achieve maximum benefits and rationalize the whole project.

3 State of the Art

3.1 Background

While we were asking ourselves how we should start designing the questionnaire, we found a lot of different kinds of recommendations and hints. From the first sight, there is a huge amount of resources, but when looking closely, we see that not most of them are, what it is called, strong recommendations. First of all, more of them are really specific and can be applied in limited number of cases, such as survey in medicine and health care [19]. Second of all, even more or less general approaches do not cover most of the design issues. Of course, it is never possible to prevent everything, but to think broadly is an option. So, some of the approaches have really good points, but if we are the ones who are designing a questionnaire, we should be able to investigate a huge amount of these approaches, to get useful information for ourselves. And it definitely requires a lot of time. Moreover, while investigating a great number of information sources, it is easy to be confused and mess them all up. Therefore, in this chapter we provide an overview of the most relevant, in our opinion, recommendations and approaches for questionnaire design.

There are few almost complete schemes for questionnaire design. They are general enough and cover common design issues. They are based on following key-principles: a questionnaire survey is only appropriate for certain research questions and the aim of the questionnaire is to obtain suitable information for statistical analysis [46, 30]. The approaches are highlighting that the validity of a questionnaire is based on clear stated research questions. Clear and in details discovered information that states what is a desired outcome of a current survey. It is the most common and general approach for any survey to design. First, we must state research questions and objectives, then, plan what kind of information we can get from a questionnaire and how it will help us to achieve research's goals.

Sounds not really difficult to do, but it is far from common use. But we will come back to it later in this chapter. One of these approaches specifies in details each questions' type. It clarifies some of the bias sources in questionnaire. It proposes to avoid bias by writing questions in short, clear meaningful manner. It suggests to avoid usage of negatively stated questions, questions where respondent is confused either agree or disagree with the statement. Moreover, it recommends to avoid double-barreled questions as well. Double-barreled questions are questions, which pursue two goals. Hence the respondent might be agree with the first part, but disagree with another.

Another significant issue of the questionnaire design, which is covered by these approaches, is a cover letter. Cover letter stated there as "*a social relationship between the respondent and the*

researcher or interviewer" [30]. The convincing and encouraging cover letter will always increase response rate for any survey [12, 33]. Recommendations provide that cover letter should include clear instructions, engaging questions, and it should be always followed by brief reminders after first invitation has been send out to the target population.

Points, given above, are common for both approaches, but unlike first one, second approach has few more advantages [28]. First, it introduces a verification part of the questionnaire design. Authors call it *a pilot work*. But by the content, this pilot work is not exactly a verification. The aim of it is to determine specification for a questionnaire (together with issues described above). But it is relatively new and different from the other approaches. By the content of specification, authors mean statement of variables, questionnaire should contain. In this approach main stress is put on questionnaire's job: *its function is measurement, and the specification should state the main variables to be measured*. The idea is behind of key statement, *questionnaire should obtain information suitable for statistical analysis*. From this point of view, current approach has been built. So, we should think about next issues, when we are designing a questionnaire according to this approach:

- How large sample would be?
- Which is the sample of population?
- Do we intend to approach the same respondents more than once?

The aim of these questions is crossing with idea behind research questions approach [30]. What is interesting in current approach, is a pilot work. Here it is addressed to design, not verify, question types. But the idea can be re-used. Pilot work suggests at first to design all multiple-choice questions as free-answer ones. Then, run this questionnaire with all open-ended questions in a small group (pilot group, around 50 people) and then, design multiple-choice answers. For the pilot work, authors also propose to divide questionnaire into short questionnaires and run each one separately in pilot groups [6]. Along with this technique, questions, which are causing some bias, can be identified and reworded. Approach also states that cover letter should be piloted as well.

So, basically, what could be done, using a pilot work, is that we run "fresh" open-ended questionnaire in a target, perhaps trusted group, and then we use the outcome to edit the questionnaire. Not just to design multiple-choice answers, but to verify and test the questionnaire.

Nevertheless, current approach has few hints different from the previous one. It suggests to start with factual questions (age, gender etc.), followed by attitudinal ones; it suggests to repeat some questions in a different contexts in a questionnaire. Definitely, it would double-verify respondent's answer, but it also makes questionnaire longer, which is the thing we should try to avoid [38].

It is clear that by using any approach, we cannot get in without good research questions. Therefore, we decided to present short overview for the research design [25]. We always start from the problem that defines our goal. Stated problem later should be divided into small sub-problems. Then, to each subproblem one or few research questions can be formulated. When we have stated research questions, we start to think which data will help us to find the answers. If such data must be collected through a survey, it is our case. It is important to stay open mind about what we may or may not discover in our data. We should just gather data, which is relevant to our research questions. And then analysis of this data will give us clear answer.

Another topic, on which approaches of questionnaire design focus, is minimization of errors [9, 11]. There are four common types of errors in the surveys:

- adequate coverage of entire population,
- sampling error,
- nonresponse, and
- measurement.

Therefore, the whole concept is based on preventions of these errors. What we personally found interesting, are "the ways of increasing the benefits of participation". That is exactly what should cover letter contain: information about the survey, need for advice, positive regard, tangible rewards, social validation etc. We found these suggestions very useful for our methodology as well.

How to design survey if instead of research questions, research is driven by hypotheses? We just need to apply few changes to the main approach [5]. Any research should be theory-driven. Same applies to the questionnaire. Each question in a questionnaire should be justified against theoretical purposes of the research. Unlike research questions, we should transform hypotheses into a list of concepts, categories. Each concept then refer to one or more variables. When this transformation is done, we have list of the variables, supporting our research. Afterwards, variables should be sorted into two groups: dependent and independent. To develop questions, the variables we already have, we need to follow few principles, which current approach states:

- do not cram more than one dimension into a single question,
- do not frame question into the negative one,
- ask specific and concrete questions rather than abstract ones,
- give clear instructions.

We can see that overall principles are the same for different approaches. That is why we called them key-principles in the beginning.

3.2 Questionnaire: What are the Advantages?

We have already briefly discussed some benefits and advantages of questionnaire. Questionnaires are quite flexible, when properly conducted, and extremely valuable tools of the research. Questionnaire obviously is a good, reasonable, commonly used tool to collect the information from the population. But what is the questionnaire's business and impact alignment in information security? How can it be useful for the information security management, except general advantages in any other field? To answer these questions, we need to take a closer look to the goals and aims of information security management.

Management of information security is addressed to maintain security inside the organization and on average level, create and support standards. Management of information security, like any other management, takes the largest part in the field, unlike the others (fig.2). In other words, it controls the whole business, called information security. Clearly, there is no doubt for how much this process is important and critical for the whole field. Hence, mistakes, made on management level, cause serious problems. Where do mistakes start? Rhetorical question. Mistakes can be made in any step, but it is extremely difficult to offload mistakes, which were made in the foundation. By foundation we mean a study, a research. Most common area of the survey research in information security is information security culture. There exists a huge amount of different frameworks and approaches of how to measure, evaluate, and discover information security culture within organization. We already have made one example in this thesis. The survey, which is chosen for the re-evaluation, is created to measure effectiveness of information security awareness program within Gjøvik University College. Basically, there are two surveys measuring information security awareness within the enterprise before and after the training had place. Another example, we want to present here, is a decision support system for the management of information security [35, 36, 37]. The system, which uses questionnaires as a tool for gathering information, is used for the following analysis and decision-making.

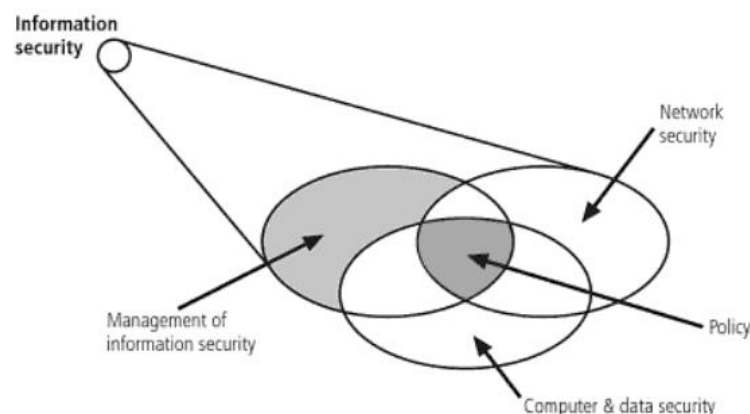


Figure 2: General components of information security [45].

We present in this chapter some of the most demonstrative examples of questionnaire's usage in terms of management of information security. Information security management in the organization focuses on employees' behavior [24]. To evaluate current behavior, surveys are frequently used [45, 22, 7]. What if we can not just evaluate current situation, but change it, using the surveys? We need to go to marketing to learn more about such technique [32, 46, 26]. In marketing, such approach is successfully used for decades.

The aim is to put in respondents' minds some new ideas about products or services. And it does not really matter, what respondents are answering. What matters is that they think about after the questionnaire was completed. In this matter, questionnaire simulates the ideas, not necessarily gathers the data. It is a covered advertisement inside the questions. For example, there is a really trivial question, which has few response options. And, let us say, the correct answer could be variable. So, question has one "obvious" answer and few more completely wrong. Respondent, of course, will choose correct one. An example of such question would be "What is the best transcontinental airlines of the year?" And for the response option, we will have just practically one transcontinental airline and the rest of the airlines, which fly within one region. We will choose the only one transcontinental, because the rest we basically even do not consider as the options. And definitely we will keep in our minds name of that chosen company. More likely, next time we will buy tickets from it.

Current example is just made up to demonstrate what can be possibly done. It might be designed not so obvious, but hidden. Nevertheless, questionnaire can be possibly used to influence respondents' opinions. So, why would we apply same thing in information security research? Why we just add few "tricky" questions to usual questionnaire, which measures information security culture of the enterprise? We can establish new concepts within the organization, using routine surveys.

Now, we would like to come back to one of the most successful, in our opinion, examples of questionnaire's usage in information security culture evaluation.

To maintain security culture in the enterprise is extremely difficult. But what is even more challenging, is trying to change it. Information security culture needs to be properly understood. Two core elements of organizational culture are basic assumptions and beliefs [35]. Information security culture is based on the same elements. Same as organizational culture, it cannot be created once and then just used. To ensure that, it corresponds with the business goals and objectives in information security, culture should be created and managed constantly. Information security culture is a management cycle. Its goal is to "sell" information security awareness to the employees, same as in marketing. Therefore, it is required to be monitored and constantly observed. Authors have proposed the tool supported management of information security culture [36]. They also claim a standardized questionnaire as the best suited instrument for a tool supported assessment. As we can see on fig. 3, each component of decision support systems contain questionnaire based survey.

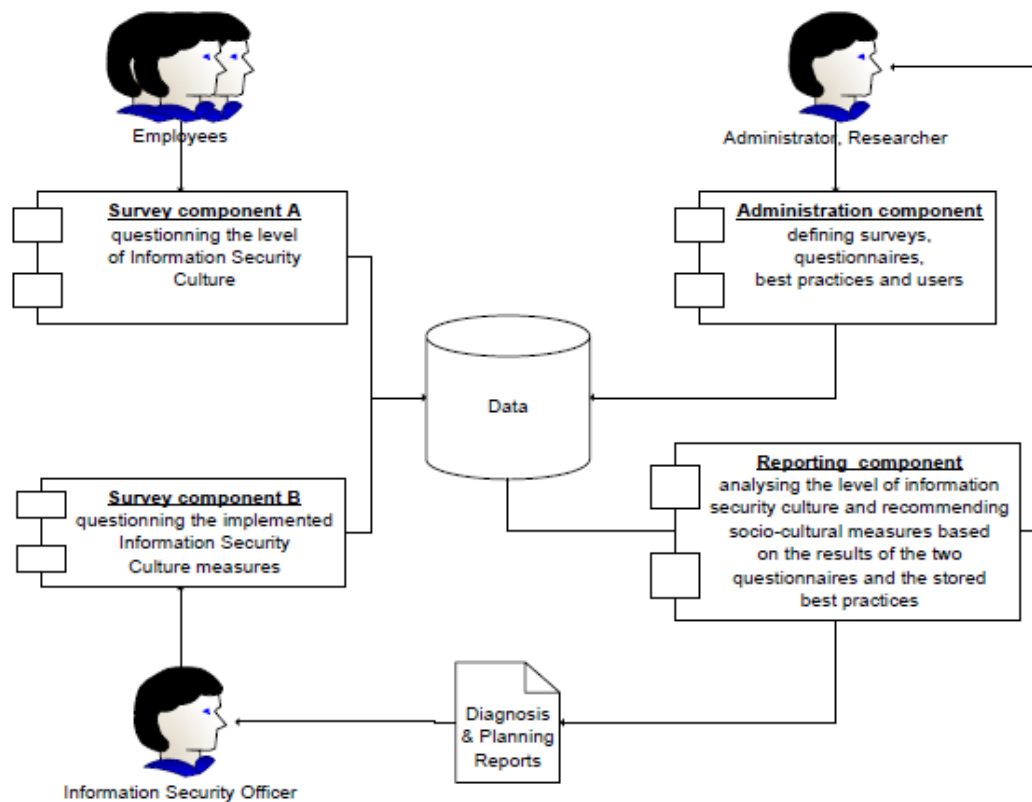


Figure 3: Architecture of the Information Security Culture Decision Support System [36].

Systems are not decision automatons, but they can help user to prepare for decision making by surveying, filtering, completing and aggregating information. What we are thinking about is that a methodology of questionnaire design would find its place in such systems. Therefore, a methodology has a future for further development and improvement, and can be used for both, business and research, purposes.

Besides information security culture, questionnaire finds its place also in many other aspects of information security management. For example, there are some empirical studies, which are examining the approach, according to which, corporations make information security expenditures decisions [17, 20]. This study gives understanding of the processes and motivations inside the top management of the business. Certainly, such survey would be a huge help to understand these decisions and make them to be made in information security interest. There are many other examples, but what connects them is the lack of proper questionnaire design. Therefore, we have decided to create a new methodology of questionnaire design within information security management field.

4 Methodology

In this chapter new methodology of questionnaire design is given. Based on related work, our experience and concept, which is described before in this report (see fig.1), methodology has been developed, explained and further tested (ch. 5).

4.1 Methodology of Questionnaire Design

Development stage of a questionnaire design is a really important part of the survey. There are a lot of things, which should be taken into account on this stage. In the state-of-the-art (ch. 3), has been mentioned that questionnaire design should always start from stated and well-understood research questions. There is no doubt that they identify the whole concept of the design. They state final goals, which survey is meant to achieve. There is no difference, if research was designed with hypotheses, not research questions. Then just the aim of a survey is to either confirm or refuse them. The point is that survey is addressed to find out something, not just gather information. Simply gathered information is useless, if it is not going to be analyzed. The results of analysis are the goal of survey. To what gathered data supposed to give an answer. This is a simple point, but crucial for questionnaire design.

When keeping this in mind, we designed an overall methodology of questionnaire design. It is unique at the point that it does not just design a questionnaire, it tests and verifies it on the development stage. This condition appeared to be novel for questionnaire design approaches.

The weak or unsatisfactory questionnaire design has an impact on the results. It reduces significance, validity, and might even confuse whole research findings. Wrongly stated questions in a questionnaire cause a lot of bias in a survey. Besides, questionnaire is only appropriate for certain research questions. Correspondence between questionnaire content and research's goals gives a validity to a survey. Moreover, except following research questions, we should keep in mind that the data we are going to collect, will be further statistically analyzed. This gives us second key-object: *it is also extremely important to design questionnaire, from which we can obtain suitable for statistical analysis data* [30]. If collected data fails to be analyzed or, at least, transformed to suitable for statistics form, survey does not make any sense. Time and resources are wasted, research goals have failed. Unfortunately, while developing a research strategy, no one spends significant time on a questionnaire design. We concentrate more on following stage, evaluation. Sometimes, we are even getting hung on results interpretation and do not notice that gathered data is far from how it supposed to be. Therefore, we present a new methodology of questionnaire design.

The conceptual idea for the methodology has been already presented in the report (fig. 1).

Now it is time to expand the concept. On fig. 4 is shown a transformed scheme of the methodology. The scheme is given as an algorithm. The input for the algorithm is a framework. From it starts the developing stage. Framework consists of a set of rules and techniques, which are addressed to help to design first-version questionnaire. **NB!** Framework is a following stage from research questions and research methods choice. It is a start-point in a methodology of questionnaire design, not whole research. Next stage is exact first-version questionnaire designing. The outcome of this stage is a "fresh" questionnaire. Afterwards, the loops are available. Here approach is getting flexible. The loop is called "updating versions". According to the name, it is clear that here questionnaire after some manipulations is going to be changed - updated. This loop on fig. 4 will be explained in details a little later in this chapter. The outcome of versions update is a complete questionnaire. It can be carried out to the target group.

We want to highlight one more time that the main body of the algorithm is versions update. it is a complex unit, which is divided into three interconnected loops. The explanation and description are following.

4.1.1 Framework

The main design principle, "KISS", is also applicable to the questionnaire. *Keep it short, simple.* Any questionnaire, we are designing, should keep the medium [40]. Keep questionnaire as short and simple as possible. Short questionnaire does not require a long time to answer it, hence rate of completed surveys will increase. Simplicity of the questions guaranties that it will be well-understood by respondents. When questions are clearly understood, more accurate data can be collected.

This is what design principle "KISS" gives us. It should be followed during the whole development process of a questionnaire. At this stage, we should already have clearly stated research questions, and questionnaire should be a justified choice to answer current questions. So, all decisions related to research questions and choice of methods are done by this stage. We mentioned above that these issues are not included in the questionnaire design, they belong to research strategy design, but they are required.

Thus, questionnaire should be as short as possible and as simple as possible. From the other hand, questionnaire's content should fully cover the research topic. To keep a balance between shortness and fully covering, following recommendations are given.

So, we start questionnaire design from identifying the information we want to gather. Research questions should help with it. When it is decided, we can write general questions, answers on which will provide us with desirable information. Each question should provide just part of a desirable information! If some of the questions give information besides desirable or it is not clear, it is always better to skip such questions at first place.

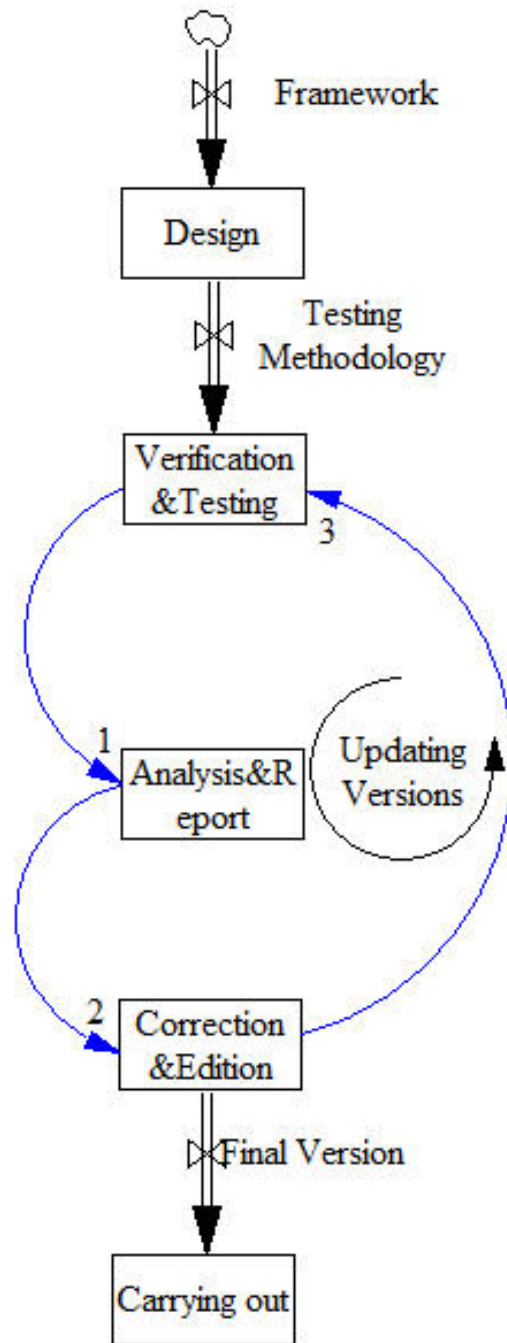


Figure 4: Complete scheme of the methodology.

After eliminating all not "perfect" questions, it is time to group them.

Questions order

We tried to keep questions, which provide just desirable information to our research. In any case, "desirable information" does not restrict us enough. Therefore, following grouping is needed.

The questions are better to place into three groups. According to which information answer on the question provides, we should divide all questions into next groups: MUST to know, USEFUL to know, and NICE to know [40]. Regarding "KISS", more likely, we have to discard last group of the questions in any case. In really rare cases, when the questionnaire is short, NICE to know group can be also included. Such grouping helps to keep questionnaire medium. Skipping last group will not influence goals of the survey, because necessary questions stay kept.

When all questions are sorted into such groups, they can be also sorted within the groups. This sorting is slightly different and might change a previous grouping. The aim is to sort questions in such manner that in the beginning "easy" questions are given and in the end, more difficult. By "easy" questions, we mean easy, pleasant to answer questions. More complicated questions should be placed closer to the end. The logic behind this sorting is simple, when respondent has already answered most of the questions, he feels pity to skip a questionnaire, when it is close to the end. Mostly, people prefer to finish it. This simple trick provides higher rate of completed surveys.

When all questions are sorted, we can go to the next step, design question types.

Question types

Choosing types of questions is really important for the questionnaire design. First, it identifies types of variables for the future statistical analysis. Second, it also influences willing of respondents to answer the questions. Therefore, types of questions should be designed carefully.

Most common types of questions, presented in questionnaires, are following: multiple-choice, numeric open end, text open end, rating scales, agreement scales, etc [8, 10]. Generally, we can categorize them into two groups: close-ended and open-ended questions. For all close-ended questions the response options are specified, and for open-ended the response is free [41, 31]. When designing a questionnaire, should be kept in mind that open-ended questions are very difficult to analyze. So, whenever is possible, open-ended questions should be avoided, unless there is no other option.

Question types should be designed according to information each question provides. Moreover, question types define with which variable we will work later. The easiest question types for following analysis, and therefore, most commonly used ones, are multiple-choice and/or rating scale questions. Close-ended questions are appropriate under few circumstances:

- The number of possible responses is small.

- The response is a specific answer.
- Researcher wants to force respondents to reply with a limited set of the categories.

From the other hand, open-ended questions are more appropriate for the exploratory research. But it is really difficult to statistically analyze this type.

To construct response options for the close-ended questions, next approach can be applied. If we are not completely sure which options should contain question, we can leave for first-version questionnaire all questions (particularly, close-ended) open. In the next stage in our methodology, testing and verification, such first-version questionnaire will be carried out to a trusted group, and analyzing the results from a trusted group, we can design response options. Moreover, we can repeat same iteration with second-version questionnaire to edit these options, and so on. Description is following.

When constructing choices for close-ended questions, order is important. For example, for rating scale question all scales have to be identified. Not just maximum and minimum scales, but all. Also, order of choices for similar close-ended questions in the questionnaire should be kept constant. If we have set of rating scale questions with scale from 1 to 5 in our questionnaire, then scales for each questions should be presented in identical order.

There is one additional trick related to the question order. When questionnaire has a set of same type questions with same response options (for example, 10 rating scale questions with same scale from 1 to 5), it is better to insert in between different questions. The point is that when we are answering different questions with completely same options for the answers, we will automatically choose same answers in a row. Unlike if we have changing response options, we will concentrate on the questions more.

4.1.2 Testing and Verification

Next part of the methodology is testing and verification. It is the most important part of the whole concept. This is a novation for the questionnaire design. The aim of it is to test and verify questionnaire on the development stage, before carrying it out to the target group.

It works as version updates. After each iteration, results will be analyzed and questionnaire will be edited. It does not just help to design response choices, it verifies questions' content. Do they truly provide desirable and expected information? Are they clear to respondents same as for designer? Do respondents understand what the question states? Does questionnaire cover the topic completely? And a lot of other questions.

Versions updates consist from two loops (fig. 1): first works through carrying each version of questionnaire out to a trusted group, second, a little more complicated, is based on semi-structured interviews in trusted group. Semi-structured interviews, used in second loop, are built on questions from the designing questionnaire.

First loop has pretty simple iterations. At the beginning, we design first-version questionnaire. It has, more likely but not necessary, all open-ended sorted questions. Then, we send this version to a trusted group. This will be a first iteration. After collecting data from a trusted group, we analyze it: eliminating some of the questions, adding another, re-sorting questions, re-wording some, perhaps, designing response choices. This is called analysis and report. Subsequently, we repeat same procedures until following iteration will stop give new results and/or outcome will be constant. After all editions and corrections, if questionnaire is finally verified, the obtained results from *different* trusted groups should be the same (not exactly same answers, but same style). First loop, unlike the second, has one unique and ponderable advantage. It allows us to measure average time, respondents need to complete the questionnaire. Time to fill in the questionnaire is really important, because if questionnaire is too long, then we can be sure that on voluntarily basis, it will not have a high response rate.

Basically, we can use just a single first loop to test and verify questionnaire. But we insist to use both loops for testing and verification.

Now, let us describe in details second loop's procedures. Unlike first one, it is addressed more on verification of content. First, we need to build a semi-structured interview based on our questionnaire [27, 28]. When an interview is constructed, we can choose few (depends on the topic, but usually 5-10 people) individuals from a trusted group. The aim of interviews is not to collect the information about the topic from interviewees. The aim is to analyze questions when sitting face-to-face with the respondent. In live communication, it is so easy to see if there are some confusing or unclear questions. Pretty often, when we are designing a question, we use some terms, which are obvious for us, but not as for the others. Interviewer can easily see such issues. Besides, if somethings is unclear, you will be definitely asked by interviewees to clarify yourself. Number of interviews is varying. As far as we can see there is no more things to change according to the interviews, we can stop.

As we can see, both loops, from one side, repeat each other, but from the other side, each has unique benefits. That is why we suggest to use both. Loops can be used either in parallel with each other or one by one. In the following chapter 5, we use second loop to verify questionnaire, then, we use first one to test improved version and design response choices.

Trusted Groups Selection

We emphasized above the usage of *different* trusted groups for iterations. In the case that we have few iterations in each loop, it is reasonable to change trusted group. For example, in a first loop, we can use one trusted group per iteration, but for the interviews we can have one trusted group around 10 individuals, which will be enough. Responses are varying for individuals, who have never seen questionnaire before and who have just filled it in. Of course, versions of questionnaire differ from each other, but the concept is constant. Because we need a fair point of view for the verification, it is recommended to change groups for each iteration in the loops.

What is exactly meant by trusted group? *Trusted group can be chosen from the target group of*

participants, taking into account such factors as the level of knowledge of the individuals in practical research and expectable manner of their behavior [46]. The number of groups depends on how many iterations we want to run. Moreover, number of individuals in a single trusted group is not high as well. For example, for the second loop we need one trusted group with around 10 individuals. For the first one, we need few, approximately same size, trusted groups. Therefore, we present a trusted group selection. Trusted group selection aims two objects: trusted groups' shifting and final-version of questionnaire validation. First object is clear. But what about validation? When should we stop to verify and test questionnaire and can finally send it out? To answer these questions, we must check if the quality of questionnaire is satisfactory, if questionnaire provides desirable outcome and does not require corrections anymore.

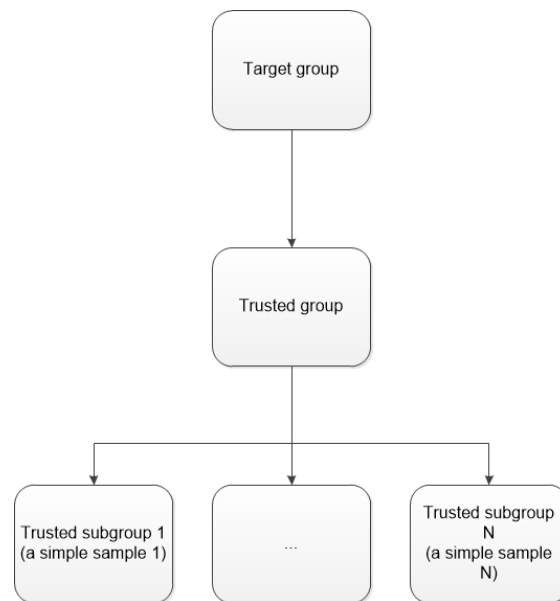


Figure 5: Trusted group sampling.

To achieve this goal, we modified sampling approach. Approach, used to sample a population, can be applied within trusted group (fig. 5). Basically, our population is a trusted group and samples are trusted subgroups. If we have two samples from the same trusted group, then responses on the same questionnaire of both subgroups should be similar. Because both subgroups are chosen from the same group of trusted individuals, therefore, responses will have approximately same distribution.

To select a sample (subgroup) of the population (trusted group, in our case), each individual in the population has to have equal chances for being selected [18]. Unlike population, individuals from a trusted group must have higher, than others, level of knowledge related to the research topic and expectable manner of behavior. Hence, the individuals from the population do not have equal chances to be selected. But when the trusted group is already selected and we need just to divide it into few small subgroups, then rules stated above apply completely. Because

we treat trusted group as a population, and subgroups, as samples of the population.

4.1.3 Cover Letter

For questionnaire to be successful, a social relationship between researcher and respondent should be established [30]. Cover letter is same important as a questionnaire. Cover letter should include a convincing and prepossessing introduction, which is accomplished by a questionnaire with clear instructions and engaging questions. Therefore, we suggest to test cover letter as well.

A cover letter should include an information like research topic, why this survey is needed, who is collecting the data, how the data will be used, anonymity of the respondent, contact information, which can be used if respondent has any difficulties or questions regarding a questionnaire, and so on.

Moreover, each invitation to participate in a survey, will be more effective if it is sent out few times. It does not mean, we should send same mail few times, it means that each survey should have few follow-ups with a brief reminder.

5 Survey Evaluation: Effectiveness of Information Security Awareness Program

The survey was conducted in 2011 at Gjøvik University College within the master thesis project by Ilirjana Veseli in collaboration with NorSIS [43]. The survey was addressed to measure the effectiveness of information security awareness program within the enterprise (Gjøvik University College). It consists of two questionnaire surveys: first measures current level of information security awareness before the training and second - after the training. The questionnaires were Internet-based and sent out to the employees of Gjøvik University College in 2011. Target group included all employees at Gjøvik University College by that time, 327 people.

The aim of this survey was to investigate which kind of the training program provides higher performance and how we can increase the information security awareness within the enterprise (University College, in our case) using such trainings.

After each survey, data was collected, analyzed and transformed to the suitable form for the following evaluation. The online survey software "Enalyzer" was used to carry survey to the participants and, afterwards, export the data to the Excel sheet [2]. Nevertheless, for the first statistical analysis, IBM SPSS Statistics has been chosen [1].

Re-evaluation is requested by NorSIS to achieve better results. The previous results were unsatisfactory because of limited time for the analysis. Moreover, NorSIS suggested to design a new approach for the questionnaire design. By putting all these together, we concluded that this survey will be a perfect example to test our methodology of questionnaire design. Evaluation of the design might and will influence the relevance of the final results.

5.1 Qualitative Analysis of the Dataset

In this section, we give information regarding to the data for the re-evaluation.

Both surveys were carried out to the participants online using "Enalyzer" software solution. "Enalyzer" is a survey software with advanced functionality [2]. This software solution was used at Gjøvik University College in 2011. "Enalyzer" makes possible to convert gathered responses to the Excel data sheets and store data in its databases. "Enalyzer" is a reasonable solution if the following statistical analysis is going to be performed in Excel. However, the statistical analysis at first place was performed using IBM SPSS Statistics [1].

The originally collected data by "Enalyzer" is inaccessible and for the re-evaluation, data in SPSS format was provided. The reasons why originally gathered data is inaccessible are few. First, Gjøvik University College refused to use "Enalyzer" software solution, therefore, software is

unaccessible (license). Second, author used her University College e-mail address to login to the profile on "Enalyzer" web page. According to the policy at Gjøvik University College the e-mail addresses of graduated students are abolished. So, the e-mail address does not exist anymore. NorSIS does not have any backups of this dataset.

The dataset was obtained from the author via e-mail. Because two years passed, since survey was done, we could get just three files from the author. To summarize, currently for the re-evaluation process we have the following files:

- "Enalyzer" report, "First Questionnaire" - Excel sheet with exported data from "Enalyzer". Data is presented in percentage graphs regarding to each question from the first questionnaire,
- *.sav file with data in SPSS format,
- second *.sav file with SPSS data from the second questionnaire.

At our disposal are also original master thesis report [43] and data files, described above. While analyzing the dataset, we faced next challenges:

- *.sav file, claimed to be first survey, includes more variables (questions) than actually first survey does.
- Names of variables given in *.sav files for the set of rating scale questions, do not correspond to the questions in surveys. Basically, instead of the actual names, all variables' labels start with the same sentence. According to the same values (scale from 1 to 5), it is impossible to identify the questions.

To deal with these issues, we contacted one more time the author of the study. All the information we could obtain from her, we have already investigated. Using the description, given by author in the thesis report, we figured it out by manipulating the data next:

- In the first survey's *.sav file, there are 6 empty rating scale questions and 2 more empty questions, which were split into 8 variables. Last 2 questions belong to the second survey, not the first one.
- Grouping of the questions, given in thesis report, does not correspond to the format of the data in both *.sav files.
- By comparing graphs, given in the report, with graphs we created using the data, we could identify the order of same-named rating scale questions in the dataset.

The time difference between when the survey was actually conducted and when it is going to be re-evaluated, makes the task even more challenging. To solve problems with the dataset, significant amount of time is spent. To get the data at a first place, we had to check number of possible sources: library at Gjøvik University College, the supervisor of the original master thesis

(Associate Professor at Gjøvik University College, Tone Hoddø Bakås), NorSIS, some of the staff, which were involved in this survey in 2011, senior adviser of Gjøvik University College, Nils Rui, who helped with data gathering in 2011, etc.

After the data was obtained, it was presented in a suitable format for statistical analysis using IBM SPSS Statistics.

To choose the right tool for statistical analysis, we had a meeting with one of the professors of Gjøvik University College, Associate Professor Frode Volden. He is a professional of statistical analysis, so his judgments played a key role in our decision. IBM SPSS Statistics was proposed by him as the best tool for the statistical analysis. Further investigation showed that IBM SPSS Statistics license is not provided by Gjøvik University College for the students' usage. Therefore, at the first place, in January 2013, open-platform analog of SPSS, PSPP, was chosen [3]. It has a limited properties and built-in tools, comparing with the original SPSS, but it is pretty similar to SPSS and requires same data format. So, the choice seemed to be reasonable.

Consequently, we had to spend time for the investigation and studying the software tools and techniques. We used literature and manuals for both SPSS and PSPP, because of their similarity [29, 21, 14, 42, 4, 1]. Eventually, Gjøvik University College provided a solution, remote server license for IBM SPSS Statistics, which is available also for students. According to these changes, it has been decided to switch back to original software, SPSS.

All these issues were, in some way, unpredictable and took additional time for finding the solutions. When all data is obtained and properly understood, and tools are chosen, we can move further, to the first step of the re-evaluation, questionnaire design analysis.

5.2 Questionnaire Design Evaluation

To evaluate questionnaires, it has been decided to use our methodology of questionnaire design. Based on fact that it is evaluation, not development, of the questionnaire, some differences are applied to our methodology.

Completed questionnaires are treated as first-version ones. Moreover, because of almost complete similarities between first and second questionnaires, they are merged into a single one. We merged both questionnaires in the way to keep and cover all questions from both surveys.

Detailed description is following.

Questionnaire Design Analysis

To measure the effectiveness of information security training at Gjøvik University College, survey, as a tool, has been chosen.

Survey consists of two questionnaires. Second questionnaire based on the first one and includes some additional and/or optional questions. It has been discussed above, we start re-

evaluation from analysis of the questionnaire design. We have already proved that weaknesses in the design influence significance and relevance of the statistical results. Therefore, questionnaire verification is required. From the other hand, if questionnaires would be properly tested and verified on the development stage, this evaluation could be avoided.

First questionnaire has been carried out to the employees before any information security awareness training program took place. First questionnaire consists of 26 questions, including one additional, open-ended, question, five multiple-choice questions, and twenty rating scale questions. The questions are grouped: first four questions require general information from the participant, the rest 22 are evaluating questions. Open-answer question, which is optional and additional, is placed on the bottom of the questionnaire.

Second questionnaire is based on the first one and, basically, repeats same questions. There are 29 questions and they are also grouped in the same way. The differences include: by the end of the section with general questions, three questions are added, two of them are considered by the author as optional ones; after rating scale questions set, one new question is added.

The comparison table 1 between first and second questionnaires is given below. Original questionnaires are given in [43], Appendixes A and B.

#	First Questionnaire		Second Questionnaire		Type
1	Q1_1	Gender	Q2_1	Gender	MC
2	Q1_2	Your age is	Q1_2	Your age is	MC
3	Q1_3	Working place	Q2_3	Working place	MC
4	Q1_4	You are currently employed as	Q2_4	You are currently employed as	MC
5			Q2_5	Which training group were you in	MC
6			Q2_6	Why you did not take part in the training	MC
7			Q2_7	Which of the following best describes the importance of this training for your job	MC*
8	Q1_5	I write down my passwords in a piece of paper near my computer	Q2_8	I write down my passwords in a piece of paper near my computer	RS
9	Q1_6	I save my passwords in my cell-phone or memory stick	Q2_9	I save my passwords in my cell-phone or memory stick	RS
10	Q1_7	I use passwords that I can easily remember so I don't have to save them	Q2_10	I use passwords that I can easily remember so I don't have to save them	RS
11	Q1_8	I don't have problem to tell my password to IT people if I am asked to	Q2_11	I don't have problem to tell my password to IT people if I am asked to	RS

#	First Questionnaire		Second Questionnaire		Type
12	Q1_9	Taking a line from a song and using the first initial from each word would be an example of a good password	Q2_12	Taking a line from a song and using the first initial from each word would be an example of a good password	RS
13	Q1_10	I use at least two different passwords. One is for working purposes, and one for private use	Q2_13	I use at least two different passwords. One is for working purposes, and one for private use	RS
14	Q1_11	I use the same password for different accounts	Q2_14	I use the same password for different accounts	RS
15	Q1_12	I don't use Password-protected screen saver in my PC at work	Q2_15	I don't use Password-protected screen saver in my PC at work	RS
16	Q1_13	I don't lock the door of my office during my working hours, even if I am away	Q2_16	I don't lock the door of my office during my working hours, even if I am away	RS
17	Q1_14	I open unexpected files or e-mail attachments or files, that I receive from unknown or known sender	Q2_17	I open unexpected files or e-mail attachments or files, that I receive from unknown or known sender	RS
18	Q1_15	I share sensitive information about my work with all my colleagues (such as information about projects, personal information about students, etc.)	Q2_18	I share sensitive information about my work with all my colleagues (such as information about projects, personal information about students, etc.)	RS
19	Q1_16	We regularly talk about how to protect sensitive information with my colleagues	Q2_19	We regularly talk about how to protect sensitive information with my colleagues	RS
20	Q1_17	Policy and regulations about information security disturbs or delays me doing my regular work	Q2_20	Policy and regulations about information security disturbs or delays me doing my regular work	RS
21	Q1_18	Only IT department is responsible for taking care of information security in GUC	Q2_21	Only IT department is responsible for taking care of information security in GUC	RS
22	Q1_19	In case when one of my colleagues is breaching the information security rules and regulations, I pretend that I am not seeing	Q2_22	In case when one of my colleagues is breaching the information security rules and regulations, I pretend that I am not seeing	RS
23	Q1_20	I put my paper documents that contain sensitive information in the recycle bin for paper	Q2_23	I put my paper documents that contain sensitive information in the recycle bin for paper	RS
24	Q1_21	I don't use shredder for discarding the documents with sensitive information	Q2_24	I don't use shredder for discarding the documents with sensitive information	RS
25	Q1_22	I save sensitive information in memory stick or external hard drive	Q2_25	I save sensitive information in memory stick or external hard drive	RS

#	First Questionnaire		Second Questionnaire		Type
26	Q1_23	I keep my desk clean from sensitive documents most of the time	Q2_26	I keep my desk clean from sensitive documents most of the time	RS
27	Q1_24	I write information about my work/research, or students in social networking sites (Facebook, twitter, MySpace)	Q2_27	I write information about my work/research, or students in social networking sites (Facebook, twitter, MySpace)	RS
28			Q2_28	I think more about information security in my everyday work after training	RS
29	Q1_25	My password is shorter than 8 characters	Q2_29	My password is shorter than 8 characters	MC
30	Q1_26	What do you think is a good password?			Open*

Table 1: Comparison of the questionnaires.

Table 1 is organized as follow:

- MC - multiple-choice question.
- RS - rating scale question.
- The asterisk (*) - optional question.
- Questions, which are common for both questionnaires, are placed in the same row in the table.
- Empty slot means that current question is unique and does not exist in another questionnaire.
- To keep general numeration between two questionnaires, first column with the aggregated number is given.

We start analysis of the questionnaires regarding to the main principles of questionnaire design. These principles are described in a framework of our methodology in chapter 4, par. 4.1.1.

Research questions of the study identify just a method of the research. "*Is it possible to show that an information security training increases the level of security awareness?*" [43]. Basically, this question identifies the method, survey, to show training's impact. But questions does not require any information about what should be inside this survey. Second research question of the study refers to the comparison of different types of training. In our opinion, research questions are not detailed enough to rely on them for the questionnaire design. So, additional questions, specific for a survey, should have been made.

For questionnaire design, author used similar study as a basis for her own [23]. It is also written in the report that questionnaire was supported by few interviews and was checked and

reviewed. But it was not enough. In the following section of this report, recoding of variables will be mentioned. Current issue is connected with question types design. More likely, this issue would be possible to avoid by following the methodology and running a questionnaire through the first loop.

We would like to show in the thesis how it could be practically done. Unfortunately, it has limited possibilities, because questionnaires are out-of-date and we do not have a task as a new questionnaire design for information security measurement. Since we did not investigate information security awareness topic, we base on general knowledge of the concept.

Questionnaire Verification and Testing

Questionnaire testing and verification include two loops: questionnaires and semi-structured interviews. We start from semi-structured interviews to focus on the questionnaire's content.

Semi-structured Interviews Semi-structured interviews, unlike questionnaire, are more flexible. Interview gives more opportunities to look through the questions. It makes possible not just to correct the questions, but also to add more questions, which were not included before, but are necessary to cover some issues. While interviewing, we can see the terms which also should be explained and clarified. For example, while asking the questions about how to treat paper documents with sensitive information, we figured out that most of the participants do not have a clear view on what sensitive information exactly is. Moreover, while interviewing, we see which questions do not make any sense at all. And so on.

In our case, we already have completed questionnaires. To apply our approach with semi-structured interviews, we merged two questionnaires into a single one: we took all questions, which are the same for both questionnaires, and added questions, which are unique for the first and second questionnaires. Also we eliminated questions such as age and gender from the list of questions for the interviews.

As a trusted group, we used employees of Gjøvik University College. Because the training has been done two years ago, it is difficult to find employees who did participate in the training and also have time for the interview. But the advantage of semi-structured interviews approach is that we do not need a big number of individuals in a trusted group. The aim of these interviews is to correct the questionnaire, not to collect the information. Currently in our case, we had seven interviews, after which we decided to stop. We stopped because following interview did not make impact of the questions anymore. Of course, testing of the questionnaire does not stop at this point, but the number of iterations for the second loop was found sufficiently enough.

The process is pretty easy. First, we build semi-structured interview's questions, which are based on the questionnaire. Second, after each conducted interview, we edit the questions. Editions are varying: we add some questions, eliminate others, split some questions into few, add more explanations and put some definitions. And what is really important, we also group and

move the questions.

sets of the questions after each conducted interview are given in Appendix A. Questions are presented in the tables and the description is given in Appendix A.

Our next step is to convert questions used for the interviews back to the questionnaire. But on this stage, we just convert questions and bring back some of them, which we eliminated in the beginning.

#	Question	Comment
1	Your gender is	
2	You are in age group	
3	Please, name the department You work	
4	Are You a full-time or part-time employee?	
5	Have You participated as a listener in IS training in 2011 at Gjøvik University College?	If YES, 3 next questions
6	Do You remember in which kind of training program You participated? If yes, please specify	
7	What was the reason why You could not participate in this training?	If answer on question 5 is NO
8	Do You actually think more about IS after You attended that training?	If answer on question 5 is YES
9	Do You think IS trainings are important for the employees?	
10	Do You use easy to remember PW?	
11	Do You keep your working PW somewhere is a written form?	If NO, skip 2 next questions
12	How do You keep it?	
13	What is the reason for You to keep it?	
14	Do You use same PW for Your private accounts and accounts at work?	
15	How many characters does Your working PW have?	
16	What is in Your opinion a good example of a PW?	
17	Please, describe generally a technique You usually use to create a new PW?	
18	What do You think will be a good approach to create a PW?	
19	Who do You think is responsible for IS at Gjøvik University College?	
20	Would You provide Your working PW to IT department if You are asked to?	
21	Does Your PC go automatically to login page if You did not use it for a while?	
22	Do You lock Your PC each time You are leaving working place?	
23	Do You lock door into Your office when You are leaving even for a short time?	
24	Do You share information related to Your work in some social networks? NOT Fronter	
25	Do You receive unexpected e-mails?	

#	Question	Comment
26	Do You open such e-mails?	
27	Do You download unknown attachments?	
28	Do You work at home sometimes?	If NO, skip next question
29	Do You bring from work some documents with sensitive information? (By "sensitive information" we mean " <i>any information for which loss, alteration, misuse or disclosure could adversely affect the interests of the University College or its administration, faculty, staff, students, applicants or relations therein</i> " [39])	If NO, skip next question
30	In which form do You bring it at home?	
31	Do You connect to remote servers from work on Your private devices to access the documents?	If NO, skip next question
32	Do You use VPN to connect to remote servers?	
33	When You need to work with paper documents with sensitive information in office, where do You keep them after You finished the work?	
34	To destroy the paper documents with sensitive information, what do You do?	
35	What do You think, IS policies delay or disturb Your regularly work?	
36	Do You ever discuss in casual conversations with Your colleagues how to protect sensitive information?	
37	Have You ever seen some of Your colleagues breaching IS rules?	
38	What were Your actions?	

Table 2: Questionnaire after first loop - semi-structured interviews.

Table 2 includes questions with some general comments. It does not include specified types of questions: rating scale, multiple-choice etc. To identify types of the questions, we are going to use first loop of our methodology, questionnaires (see fig.1).

We want to emphasize one more time that all the editions and changes are based on originally designed questionnaires. The content of the questionnaires was changed according to only information from interviews. We did not do any investigation of questionnaire's topic - information security awareness. The knowledge we have, related to this topic, is a general knowledge, all master students of information security at Gjøvik University College have.

Questionnaires It has been decided to run the questionnaire after interviews in a trusted group to verify it again and to design question types. The trusted group is different from the trusted group, used for the interviews. Current trusted group consists from information security master students. The decision based on few reasons: first, to access the employees one more time for pretty much the same survey is impolite, second, to select another trusted group from the employees is really difficult, because whole population is not big enough and considering response rate, chances to select it are very low. Otherwise, master students of information security are knowledgeable, more or less free to participate in such kind of surveys and easily

accessible. Moreover, trusted group for this task perfectly corresponds the requirements stated in a definition of a trusted group. Definitely, level of knowledge of information security masters corresponds desired level, and individuals of the trusted group have expectable manner of their behavior.

To carry questionnaire out to the trusted group, Google Drive tools were used [16]. Form is filled in with questions from table 2 with all questions are open-ended. 24 responses are collected. Based on the responses, types of questions are designed.

As it has been mentioned before in chapter 4, question types define measurement values for further statistical analysis. Therefore when designing question types, we should think how statistical analysis is going to be performed. So, it has been decided to transfer majority of questions back to rating scale ones. Also, the content of questionnaire has been slightly changed, to "pretend" it is a survey to evaluate information security awareness after the training program. Also, recoding issues is taken into account. Questions, identified by us as NICE to know group, are eliminated.

Q2_1 Your gender is:

- (a) Male
- (b) Female

Q2_2 You are in an age group:

- (a) Less than 30
- (b) 31-50
- (c) Over 51

Q2_3 Please, name the department You work:

- (a) Administration
- (b) Faculty of Health, Care and Nursing
- (c) Faculty of Technology, Economy and Management

Q2_4 Are You a full-time employee?

- (a) Yes
- (b) No

Q2_5 Have You participated in information security training at Gjøvik University College?

- (a) Yes (please, skip following question)
- (b) No (please, skip question Q2_8)

Q2_6 What was the reason You could not participate in a training?

- (a) I was not informed
- (b) I did not have time
- (c) It is not important for me
- (d) I do not need such training, I am an expert in this field
- (e) Other

Q2_7 Which training program did You attend?

- (a) Classroom-training (Eureka 2/3)
- (b) Discussion-training (A128)
- (c) Web-based training (e-mail with lessons)
- (d) I do not remember

Q2_8 Do You think information security trainings are important for the employees at Gjøvik University College?

- (a) There is no particular need
- (b) I had one, when I was employed
- (c) Good to have at least one per year
- (d) Extremely important

Q2_9 I use passwords, which I cannot easy remember

- 1. totally agree
- 2. agree
- 3. not sure
- 4. disagree
- 5. totally disagree

Q2_10 I keep my password from work in a written form

- 1. totally agree
- 2. agree
- 3. not sure
- 4. disagree
- 5. totally disagree

Q2_11 I keep it:

- (a) on a sticker on my desk
- (b) in a text file on my computer

- (c) saved in my cellphone
- (d) saved in my e-mail
- (e) other
- (f) I do not have my password in a written form

Q2_12 What is the reason for You to keep it like this?

- (a) It is too complicated, I cannot remember it
- (b) It changes too often
- (c) Just in case if I will forget it once
- (d) I do not have my password in a written form

Q2_13 I use same password for work and private accounts

1. totally agree
2. agree
3. not sure
4. disagree
5. totally disagree

Q2_14 How many characters does Your usual working password have?

- (a) Less than 8
- (b) 8
- (c) More than 8

Q2_15 Which one, in Your opinion, is a good example of the password:

- (a) sequence of random numbers
- (b) sequence of random numbers and letters
- (c) sequence of random numbers, capital and lowercase letters+special characters
- (d) sequence of random numbers, capital and lowercase letters+special characters, which make sense for me, but not others

Q2_16 Who do You think is responsible for information security in Gjøvik University College?

- (a) I have no idea
- (b) No one
- (c) Administration
- (d) IT department
- (e) Everyone who works or studies at college

Q2_17 I feel absolutely secure to share my password with people from IT department

1. totally agree
2. agree
3. not sure
4. disagree
5. totally disagree

Q2_18 My PC does not go automatically to login page when I am not using it for a while

1. totally agree
2. agree
3. not sure
4. disagree
5. totally disagree

Q2_19 I do not lock my office when I am leaving it for a while

1. totally agree
2. agree
3. not sure
4. disagree
5. totally disagree

Q2_20 I share information related to my work via social networks (not Fronter!)

1. totally agree
2. agree
3. not sure
4. disagree
5. totally disagree

Q2_21 I open unexpected e-mails

1. totally agree
2. agree
3. not sure
4. disagree
5. totally disagree

Q2_22 I download unknown attachments

1. totally agree
2. agree
3. not sure
4. disagree
5. totally disagree

Q2_23 Do You work at home sometimes?

- (a) Yes, sometimes
- (b) No, never

Q2_24 Do You bring home documents with sensitive information from work? By "sensitive information" is meant "*any information for which loss, alteration, misuse or disclosure could adversely affect the interests of the University College or its administration, faculty, staff, students, applicants or relations therein.*"

- (a) Yes, if needed
- (b) No, I never work at home with such documents (please, skip next question)

Q2_25 How do You bring them home?

- (a) I save them in cloud services (Dropbox, Google Drive etc.)
- (b) I copy them on external devices (memory stick, external HDD, CD/DVD etc.)
- (c) I print them out
- (d) I connect to remote server
- (e) I have them on my working laptop

Q2_26 To connect to remote server, I never use VPN (virtual private network)

1. totally agree
2. agree
3. not sure
4. disagree
5. totally disagree

Q2_27 I keep paper documents with sensitive information on my desk after I finished working with them

1. totally agree
2. agree
3. not sure
4. disagree

5. totally disagree

Q2_28 I put paper documents with sensitive information in my recycle bin in office

1. totally agree
2. agree
3. not sure
4. disagree
5. totally disagree

Q2_29 Information security policy and regulations disturb my routine work

1. totally agree
2. agree
3. not sure
4. disagree
5. totally disagree

Q2_30 If I see some of my colleagues are breaching policy information security rules, I pretend I do not see

1. totally agree
2. agree
3. not sure
4. disagree
5. totally disagree

Second questionnaire is re-designed. Number of questions is lower, different types of questions are mixed in a questionnaire, same coding is saved among the questions, NICE to know questions are eliminated, more explanations are given.

5.3 Statistical Analysis and Results Interpretation

The idea to measure effectiveness of the training program was to compare levels of awareness of the employees before and after the training. Levels of awareness were measured by conducting two surveys and additional interviews in a focus group ([43], ch. 2, par. 2.7.2). The data collected by interviews, were not presented for the re-evaluation. Both surveys have similar set of questions and were carried out to a target group before and after the training. The target group included all employees of Gjøvik University College in 2011 (327 people). The response rate of the first survey, table 6, is 159 responses (48,62%). For the second survey, table 7, response rate consists from 110 completed questionnaires (33,63%) ([43], ch. 5). First survey was filled in by respondents before any information security awareness trainings. Second one was filled in by both types of respondents: who still did not participate in any information security awareness

trainings and who did in one of them. There were three types of training conducted at Gjøvik University College: classroom-training, discussion-training, and web-based training ([43], ch. 4, par. 4.5).

Sent survey requests	327
Completed Surveys	159
Completion Rate	48.62%
Uncompleted forms (after started)	10
Average Completion time	9 minutes

Figure 6: Responses, first survey [43].

Sent survey requests	327
Completed Surveys	110
Completion Rate	33.63%
Uncompleted forms (after started)	7
Average Completion time	9 minutes

Figure 7: Responses, second survey [43].

The aim of re-evaluation of this survey is to find out either trainings had an impact on information security awareness level of the employees and if it did, then which program had a higher performance. To answer these question, we use obtained datasets to perform statistical analysis. The idea is to compare responses before and after the trainings (responses from participants and non-participants of the trainings) between each other and see either the significant difference exists among them. Same idea is used to compare three training programs between each other. By comparing the programs, we can figure out, do they actually have different performances.

The description of datasets is given in the section above 5.1. First, we start analysis from editing the data for the following evaluation. As it was mentioned above, the dataset from first survey seemed to be corrupted. We had to check each variable in a dataset for a match with the data in the original report. Because of lack of dataset description, we had to apply such approach. The results are following: six empty variables, claimed to be rating scale questions, are deleted; eight empty variables (multiple-choice) are deleted as well. The last eight variables are two additional questions from the second survey, which were split in a dataset into four variables each. All deleted variables from the first dataset are empty and do not belong to the first survey. By completing this eliminations, dataset is presented as it should have been. It has equal number of questions with the first questionnaire, does not have any empty variables, all variables correspond questions. The set of rating scale questions is originally presented in a dataset in a manner, when it is impossible to identify single variable to a single question (all variables have same names and answer options). By comparing made bar charts from each variable with the bar charts given in the report, correct order has been found. Therefore, all variables (rating scale questions) from both datasets are renamed, so it is possible to refer each variable to actual question in the surveys. Numbering of variables has been changed as well to the same style as in the

questionnaires. Qn_*number of question* - where n is a number of current survey (see table1).

More variables are deleted from both datasets, but for a different reason. After questionnaire design evaluation, we assume that questions Q1_9, Q1_10 and Q1_11 (respectively same questions from the second survey) are not relevant for the analysis. Question Q1_9 states that "*taking a line from a song and using the first initials from each word would be an example of a good password*" and has five rating scales: totally disagree, disagree, not sure, agree, and totally agree. Current approach for creating a password is neither good or bad. It is definitely not the worst, like taking date of birth, which is commonly used. From another hand, it is far from good, because it contains just letters and might be already in databases for dictionary attacks. The point is current question does not show any specific knowledge, but causes a bias to the results. Question Q1_10 is unacceptable for the analysis as well. Everyone uses different passwords, not even on purpose. A lot of the web-services have limitations on the password length, characters etc. Some of them force users to have long and secure passwords, some - just a number. So the question is insignificant. Q1_11 in some case repeats previous question. For the same reason, it is eliminated from further analysis as well.

On the next step, rating scale questions' values are recoded. All rating scale questions have the same scale: 1 - totally disagree, 2 - disagree, 3 - not sure, 4 - agree, and 5 - totally agree. Majority of the rating scale questions in both questionnaires stated in the way that total disagreement would be considered as the highest level of awareness. Therefore, recoding is needed. Recoding is made to the opposite scale: 1 - totally agree, 2 - agree, 3 - not sure, 4 - disagree, and 5 - totally disagree. It gives us less aware choice's score lower and more aware choice's score higher. For example, scores 4 and 5 will be considered as high aware choices, unlike 1, 2 and 3, which will be less aware choices. To keep same style for all rating scale questions (higher score - more aware, lower score - less aware), not all of them are recoded. Those, which are recoded, are given below (table 3).

We would like to add at this point, that issue with recoding variables, could be managed on the development stage of a questionnaire.

Table 3: Recoded questions from both questionnaires.

# First Survey	# Second Survey
Q1_5	Q2_8
Q1_6	Q2_9
Q1_8	Q2_11
Q1_12	Q2_15
Q1_13	Q2_16
Q1_14	Q2_17
Q1_15	Q2_18
Q1_17	Q2_20
Q1_18	Q2_21
Q1_19	Q2_22
Q1_20	Q2_23
Q1_21	Q2_24
Q1_22	Q2_25
Q1_24	Q2_27

Not all of the questions will be analyzed. According to the questionnaire design, we can identify groups of the questions (MUST, USEFUL and NICE to know). These groups should have been identified on the development stage of the questionnaire. Because we work with already completed survey, we can use these grouping to chose right questions for the evaluation's goal, measurement of the effectiveness of information security awareness program. To the group MUST to know questions, which are exactly evaluating level of awareness, should be included. Our MUST to know group is a group of questions, comparing whose responses, we can measure the awareness level. So hence, we can compare those levels between participants and non-participants afterwards. USEFUL to know group can be specified based on such information like level of awareness among the departments of Gjøvik University College. Or level of awareness regarding gender, age groups etc. This group does not related to the aim of this evaluation, therefore, is skipped from the analysis. Same thoughts are applied to the third group, NICE to know. Questions, detected to belong to the last two groups, do not make an impact on the evaluation's goal. Following questions are eliminated for further statistical analysis, for both surveys respectively: Q1_16, Q1_17, Q1_18, Q1_19. We understand, that it is doubtable decision to make, and author of the survey might be disagree. But we are sure it will not influence main results of the analysis.

The detailed response data analysis is given in original report, chapter 5 [43]. Because we did not collect the data and anyhow cannot influence the response results, we refer to the original report to avoid repetition.

Remained questions, which are chosen for the further statistical analysis, are listed in table4. The questions are also grouped into two sets to simplify the analysis. Unlike original evaluation, it has been decided to not merge variables in a dataset according to this grouping. Because of different number of questions in both groups, common range will be different. So, instead of range from 1 to 5 for majority of variables (the only exception is Q1_25/Q2_29), for merged variables, range will depend on number of included in merger variables. It might also influence

the perception of following plots and numbers in the analysis.

Table 4: Grouping.

Group#	# First Survey	# Second Survey
1: Password management	Q1_5, Q1_6, Q1_7, Q1_8, Q1_12, Q1_25	Q2_8, Q2_9, Q2_10, Q2_11, Q2_15, Q2_29
2: Policy regulations awareness	Q1_13, Q1_14, Q1_15, Q1_20, Q1_21, Q1_22, Q1_23, Q1_24	Q2_16, Q2_17, Q2_18, Q2_23, Q2_24, Q2_25, Q2_26, Q2_27

The grouping divides analysis into subtasks and makes easier results' interpretation. Grouping is based on the questions' content and is overall. In the original analysis, questions were grouped as well. Questions were designed at first place to cover specific topics, groups. But, for some reasons, grouping was not kept in the questionnaires. Because we did not design current survey and training programs, our grouping does not pursue same goal as in original evaluation.

All preparations for statistical analysis are finished. First, we want to present frequencies tables 8, 9, 10, and 11. In the tables can be found mean, standard deviation, variance, skewness, range, maximum and minimum for each question. We are interested in measures of central tendency, variability, and percentiles. Mean, skewness, and range are used for the descriptive comparison between first and second surveys. Bar charts with percentage distribution are given in Appendix B. Bar charts, for each question from both surveys respectively, are placed one by one. The percentages are given on each bar.

As it was said before, higher score corresponds to higher level of awareness. By simply looking into content of tables 8 and 9, we can see that means do not really differ for surveys one and two in the first group. Skewness measures to what extent a distribution of values deviated from symmetry around the mean [44, 15]. Negative skewness in tables shows that distribution of values moved to the right from the mean, closer to higher score. Which basically shows us positive tendency. Although slight differences between means and skewness, this analysis is not capable to tell us anything about significance of results.

Statistics

		I write down my passwords in a piece of paper near my computer	I save my passwords in my cell-phone or memory stick	I use passwords that I can easily remember so I don't have to save them	I don't have problem to tell my password to IT people if I am asked to	I don't use Password-protected screen saver in my PC at work	My password is shorter than 8 characters:
N	Valid	159	159	159	159	159	159
	Missing	0	0	0	0	0	0
Mean		4.37	4.58	3.50	3.975	3.04	2.40
Std. Deviation		1.156	.950	1.427	1.1957	1.655	.585
Variance		1.336	.903	2.036	1.430	2.739	.342
Skewness		-1.760	-2.445	-.688	-.873	.031	-.356
Std. Error of Skewness		.192	.192	.192	.192	.192	.192
Range		4	4	4	4.0	4	2
Minimum		1	1	1	1.0	1	1
Maximum		5	5	5	5.0	5	3

Figure 8: Frequencies - group #1, first survey.

Statistics

		I write down my passwords in a piece of paper near my computer	I save my passwords in my cell-phone or memory stick	I use passwords that I can easily remember so I don't have to save them	I don't have problem to tell my password to IT people if I am asked to	I don't use Password-protected screen saver in my PC at work	I think more about information security in my everyday work after training
N	Valid	110	110	110	110	110	110
	Missing	0	0	0	0	0	0
Mean		4.28	4.59	3.63	4.46	3.52	3.13
Std. Deviation		1.166	.912	1.326	1.046	1.561	1.015
Variance		1.360	.831	1.759	1.095	2.435	1.030
Skewness		-1.560	-2.199	-.773	-1.908	-.514	-.314
Std. Error of Skewness		.230	.230	.230	.230	.230	.230
Range		4	3	4	4	4	4
Minimum		1	2	1	1	1	1
Maximum		5	5	5	5	5	5

Figure 9: Frequencies - group #1, second survey.

Second group of questions is policy regulations awareness. We named this group based on opinion that these issues should be stated in the policy. Questions from second group describe opposite from "aware" and responsible behavior of the employees.

Frequencies for second group are given in tables 10 and 11 for first and second surveys, respectively. We can observe changes in means among group in both surveys. Basically, we can see same situation as in first group. There are slight differences for means and skewness between two surveys. But to judge for how much two surveys differ from each other, further analysis is

needed [25].

Statistics

		I don't lock the door of my office during my working hours, even if I am away	I open unexpected files or e-mail attachments or files, that I receive from unknown or known sender	I share sensitive information about my work with all my colleagues (such as projects, personal information about students, etc.)	I put my paper documents that contain sensitive information in the recycle bin for paper	I don't use shredder for discarding the documents with sensitive information	I save sensitive information in memory stick or external hard drive	I keep my desk clean from sensitive documents most of the time	I write information about my work/research, or students in social networking sites (Facebook, twitter, MySpace)
N	Valid	159	159	159	159	159	159	159	159
	Missing	0	0	0	0	0	0	0	0
Mean		4.03	4.37	4.30	4.26	4.18	3.65	3.85	4.60
Std. Deviation		1.265	.971	.940	1.104	1.174	1.258	1.074	.935
Variance		1.600	.944	.883	1.218	1.378	1.582	1.154	.874
Skewness		-1.067	-1.770	-1.381	-1.414	-1.360	-.519	-.935	-2.802
Std. Error of Skewness		.192	.192	.192	.192	.192	.192	.192	.192
Range		4	4	4	4	4	4	4	4
Minimum		1	1	1	1	1	1	1	1
Maximum		5	5	5	5	5	5	5	5

Figure 10: Frequencies - group #2, first survey.

Statistics

		I don't lock the door of my office during my working hours, even if I am away	I open unexpected files or e-mail attachments or files, that I receive from unknown or known sender	I share sensitive information about my work with all my colleagues (such as projects, personal information about students, etc.)	I put my paper documents that contain sensitive information in the recycle bin for paper	I don't use shredder for discarding the documents with sensitive information	I save sensitive information in memory stick or external hard drive	I keep my desk clean from sensitive documents most of the time	I write information about my work/research, or students in social networking sites (Facebook, twitter, MySpace)
N	Valid	110	110	110	110	110	110	110	110
	Missing	0	0	0	0	0	0	0	0
Mean		4.25	4.65	4.36	4.27	4.25	3.73	3.83	4.83
Std. Deviation		1.135	.724	.896	1.048	1.175	1.256	1.132	.539
Variance		1.288	.525	.802	1.099	1.380	1.576	1.282	.291
Skewness		-1.534	-2.442	-1.413	-1.494	-1.426	-.572	-1.005	-3.691
Std. Error of Skewness		.230	.230	.230	.230	.230	.230	.230	.230
Range		4	4	3	4	4	4	4	3
Minimum		1	1	2	1	1	1	1	2
Maximum		5	5	5	5	5	5	5	5

Figure 11: Frequencies - group #2, second survey.

To say either scores of participants of the training are significantly differ from scores of non-

participants, we need to compare means for both options. We decide to run One-Way ANOVA test, because of the next additional options for comparison: descriptives, homogeneity-of-variance selection, and multiple comparisons [34]. One-Way ANOVA is applied for second survey dataset. This dataset includes both, training-participants and non-participants, kinds of responses. Moreover, in this dataset we are also able to compare three training programs between each other. Another reason to run this comparison in a dataset from the second survey is a distribution of responses (fig. 12). Numbers of responses from each group (classroom, discussion, web-based trainings, and "None") are more or less covered by normal distribution. Unlike, if we would compare training-participants scores with non-participants scores, taken from the first survey. Even total number of responses from second survey is lower than number of responses from first one (tables 6 and 7).

Descriptives (tables 13 and 17) provide number, mean, standard deviation, standard error, and range for each level. By level, training group is meant. Also "None" as a level is included, which corresponds to non-participants' responses. The homogeneity-of-variance selection provides information about suitability of our variables for the analysis (tables 14 and 18). Significance value of Levene's test for homogeneity of variance indicates either scores for each level differ significantly [15]. To differ significantly, significance value should be less than 0.05 (for one-tailed sig. value). Mean plots for each question for both groups are presented in Appendix C. On each plot, means for all levels are given. ANOVA tables (15 and 19) provide information about existing difference between and within groups (levels). For ANOVA, two-tailed significance value is used, so the limit, under which scores differ significantly, is 0.1. If significance value lies between 0.05 and 0.1 for one-tailed value, scores can be considered as marginally different. For a two-tailed significance value, limits must be multiplied by 2. Multiple comparisons, LSD, (tables 16 and 20) are applied to compare each group (level) with others one by one. The asterisk (*) in the tables indicates that there is one pair of groups whose means differ significantly from each other (at level of significance 0.05).

Overall ANOVA results might be different from pairwise comparisons in association with different procedures used for comparisons. Overall ANOVA compares all values simultaneously, when LSD is a series of independent *t* tests ([15], ch. 12). So, LSD is our double-check of ANOVA results.

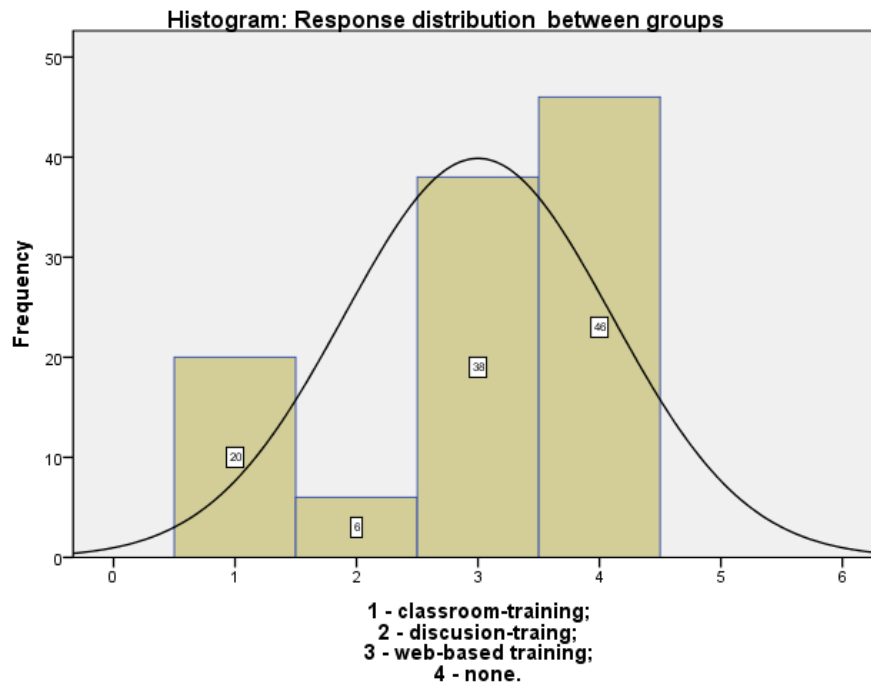


Figure 12: Second survey: percentage between groups of respondents.

Descriptives

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum	
					Lower Bound	Upper Bound			
I write down my passwords in a piece of paper near my computer	a) Classroom-training (Eureka 2/3).	20	4.30	1.380	.309	3.65	4.95	1	5
	b) Discussion-training (A128).	6	4.00	1.095	.447	2.85	5.15	2	5
	c) Web-based training (e-mail with lessons).	38	4.47	1.109	.180	4.11	4.84	1	5
	d) None.	46	4.15	1.135	.167	3.82	4.49	1	5
	Total	110	4.28	1.166	.111	4.06	4.50	1	5
I save my passwords in my cell-phone or memory stick	a) Classroom-training (Eureka 2/3).	20	4.80	.696	.156	4.47	5.13	2	5
	b) Discussion-training (A128).	6	5.00	.000	.000	5.00	5.00	5	5
	c) Web-based training (e-mail with lessons).	38	4.55	.950	.154	4.24	4.86	2	5
	d) None.	46	4.48	1.005	.148	4.18	4.78	2	5
	Total	110	4.59	.912	.087	4.42	4.76	2	5
I use passwords that I can easily remember so I don't have to save them	a) Classroom-training (Eureka 2/3).	20	3.85	1.348	.302	3.22	4.48	1	5
	b) Discussion-training (A128).	6	3.17	1.472	.601	1.62	4.71	1	5
	c) Web-based training (e-mail with lessons).	38	3.47	1.224	.199	3.07	3.88	1	5
	d) None.	46	3.72	1.393	.205	3.30	4.13	1	5
	Total	110	3.63	1.326	.126	3.38	3.88	1	5
I don't have problem to tell my password to IT people if I am asked to	a) Classroom-training (Eureka 2/3).	20	4.55	1.146	.256	4.01	5.09	1	5
	b) Discussion-training (A128).	6	3.67	1.633	.667	1.95	5.38	1	5
	c) Web-based training (e-mail with lessons).	38	4.71	.694	.113	4.48	4.94	2	5
	d) None.	46	4.33	1.117	.165	3.99	4.66	1	5
	Total	110	4.46	1.046	.100	4.27	4.66	1	5
I don't use Password-protected screen saver in my PC at work	a) Classroom-training (Eureka 2/3).	20	3.65	1.599	.357	2.90	4.40	1	5
	b) Discussion-training (A128).	6	4.00	1.265	.516	2.67	5.33	2	5
	c) Web-based training (e-mail with lessons).	38	3.53	1.640	.266	2.99	4.07	1	5
	d) None.	46	3.39	1.542	.227	2.93	3.85	1	5
	Total	110	3.52	1.561	.149	3.22	3.81	1	5
My password is shorter than 8 characters:	a) Classroom-training (Eureka 2/3).	20	2.45	.510	.114	2.21	2.69	2	3
	b) Discussion-training (A128).	6	2.67	.516	.211	2.12	3.21	2	3
	c) Web-based training (e-mail with lessons).	38	2.37	.541	.088	2.19	2.55	1	3
	d) None.	46	2.41	.541	.080	2.25	2.57	1	3
	Total	110	2.42	.531	.051	2.32	2.52	1	3

Test of Homogeneity of Variances

	Levene Statistic	df1	df2	Sig.
I write down my passwords in a piece of paper near my computer	.628	3	106	.599
I save my passwords in my cell-phone or memory stick	4.250	3	106	.007
I use passwords that I can easily remember so I don't have to save them	.499	3	106	.684
I don't have problem to tell my password to IT people if I am asked to	4.987	3	106	.003
I don't use Password-protected screen saver in my PC at work	.876	3	106	.456
My password is shorter than 8 characters:	.308	3	106	.820

Figure 14: Test of Homogeneity of Variances - questions group #1.

ANOVA

		Sum of Squares	df	Mean Square	F	Sig.
I write down my passwords in a piece of paper near my computer	Between Groups	2.655	3	.885	.644	.588
	Within Groups	145.608	106	1.374		
	Total	148.264	109			
I save my passwords in my cell-phone or memory stick	Between Groups	2.518	3	.839	1.010	.391
	Within Groups	88.073	106	.831		
	Total	90.591	109			
I use passwords that I can easily remember so I don't have to save them	Between Groups	3.535	3	1.178	.664	.576
	Within Groups	188.183	106	1.775		
	Total	191.718	109			
I don't have problem to tell my password to IT people if I am asked to	Between Groups	7.147	3	2.382	2.250	.087
	Within Groups	112.208	106	1.059		
	Total	119.355	109			
I don't use Password-protected screen saver in my PC at work	Between Groups	2.483	3	.828	.334	.801
	Within Groups	262.980	106	2.481		
	Total	265.464	109			
My password is shorter than 8 characters:	Between Groups	.486	3	.162	.567	.638
	Within Groups	30.278	106	.286		
	Total	30.764	109			

Figure 15: ANOVA table - questions group #1.

Multiple Comparisons

LSD

Dependent Variable	(I) Which training group where you in:		Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
I write down my passwords in a piece of paper near my computer	a) Classroom-training (Eureka 2/3).	b) Discussion-training (A128).	.300	.546	.584	-0.78	1.38
		c) Web-based training (e-mail with lessons).	-.174	.324	.593	-.82	.47
	b) Discussion-training (A128).	d) None.	.148	.314	.639	-.47	.77
		c) Web-based training (e-mail with lessons).	-.474	.515	.360	-1.49	.55
I save my passwords in my cell-phone or memory stick	c) Web-based training (e-mail with lessons).	d) None.	-.152	.509	.765	-1.16	.86
		d) None.	.322	.257	.214	-.19	.83
	a) Classroom-training (Eureka 2/3).	b) Discussion-training (A128).	-.200	.424	.638	-1.04	.64
		c) Web-based training (e-mail with lessons).	.247	.252	.328	-.25	.75
I use passwords that I can easily remember so I don't have to save them	b) Discussion-training (A128).	d) None.	.322	.244	.190	-.16	.81
		c) Web-based training (e-mail with lessons).	.447	.400	.266	-.35	1.24
	c) Web-based training (e-mail with lessons).	d) None.	.522	.396	.190	-.26	1.31
		d) None.	.074	.200	.710	-.32	.47
I don't have problem to tell my password to IT people if I am asked to	a) Classroom-training (Eureka 2/3).	b) Discussion-training (A128).	.683	.620	.273	-.55	1.91
		c) Web-based training (e-mail with lessons).	.376	.368	.309	-.35	1.11
	b) Discussion-training (A128).	d) None.	.133	.357	.711	-.57	.84
		c) Web-based training (e-mail with lessons).	-.307	.585	.601	-1.47	.85
I don't use Password-protected screen saver in my PC at work	c) Web-based training (e-mail with lessons).	d) None.	-.551	.578	.343	-1.70	.60
		d) None.	-.244	.292	.406	-.82	.34
	a) Classroom-training (Eureka 2/3).	b) Discussion-training (A128).	.883	.479	.068	-.07	1.83
		c) Web-based training (e-mail with lessons).	-.161	.284	.573	-.72	.40
My password is shorter than 8 characters:	b) Discussion-training (A128).	d) None.	.224	.276	.418	-.32	.77
		c) Web-based training (e-mail with lessons).	-1.044	.452	.023	-1.94	-.15
	c) Web-based training (e-mail with lessons).	d) None.	-.659	.447	.143	-1.54	.23
		d) None.	.384	.226	.091	-.06	.83
I don't use Password-protected screen saver in my PC at work	a) Classroom-training (Eureka 2/3).	b) Discussion-training (A128).	-.350	.733	.634	-1.80	1.10
		c) Web-based training (e-mail with lessons).	.124	.435	.777	-.74	.99
	b) Discussion-training (A128).	d) None.	.259	.422	.541	-.58	1.10
		c) Web-based training (e-mail with lessons).	.474	.692	.495	-.90	1.85
My password is shorter than 8 characters:	c) Web-based training (e-mail with lessons).	d) None.	.609	.684	.375	-.75	1.96
		d) None.	.135	.345	.697	-.55	.82
	a) Classroom-training (Eureka 2/3).	b) Discussion-training (A128).	-.217	.249	.386	-.71	.28
		c) Web-based training (e-mail with lessons).	.082	.148	.582	-.21	.37
My password is shorter than 8 characters:	b) Discussion-training (A128).	d) None.	.037	.143	.797	-.25	.32
		c) Web-based training (e-mail with lessons).	.298	.235	.207	-.17	.76
	c) Web-based training (e-mail with lessons).	d) None.	.254	.232	.277	-.21	.71
		d) None.	-.045	.117	.704	-.28	.19

*. The mean difference is significant at the 0.05 level.

Figure 16: LSD - questions group #1.

The overall description of applied procedures is given, let us than discuss the results. For the first group of questions, descriptives show general information regarding each level. Test of ho-

mogeneity indicates just two questions (Q2_9 and Q2_11), whose level of significance is less than 0.05. It means that just these two questions' scores differ each level significantly. Overall ANOVA identifies a single question, Q2_11, with significance level less than 0.1 (one-tailed value). This question has scores significantly different among all four levels (groups). LSD proofs that current questions' scores are different for discussion and web-based trainings.

Descriptives									
	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum	
					Lower Bound	Upper Bound			
I don't lock the door of my office during my working hours, even if I am away	a) Classroom-training (Eureka 2/3).	20	4.45	1.099	.246	3.94	4.96	1	5
	b) Discussion-training (A128).	6	4.67	.516	.211	4.12	5.21	4	5
	c) Web-based training (e-mail with lessons).	38	4.37	1.051	.170	4.02	4.71	2	5
	d) None.	46	4.00	1.247	.184	3.63	4.37	1	5
	Total	110	4.25	1.135	.108	4.03	4.46	1	5
I open unexpected files or e-mail attachments or files, that I receive from unknown or known sender	a) Classroom-training (Eureka 2/3).	20	4.70	.801	.179	4.32	5.08	2	5
	b) Discussion-training (A128).	6	4.83	.408	.167	4.40	5.26	4	5
	c) Web-based training (e-mail with lessons).	38	4.58	.793	.129	4.32	4.84	1	5
	d) None.	46	4.65	.674	.099	4.45	4.85	3	5
	Total	110	4.65	.724	.069	4.51	4.78	1	5
I share sensitive information about my work with all my colleagues (such as information about projects, personal information about students, etc.)	a) Classroom-training (Eureka 2/3).	20	4.35	.988	.221	3.89	4.81	2	5
	b) Discussion-training (A128).	6	4.50	.837	.342	3.62	5.38	3	5
	c) Web-based training (e-mail with lessons).	38	4.50	.762	.124	4.25	4.75	2	5
	d) None.	46	4.24	.970	.143	3.95	4.53	2	5
	Total	110	4.36	.896	.085	4.19	4.53	2	5
I put my paper documents that contain sensitive information in the recycle bin for paper	a) Classroom-training (Eureka 2/3).	20	4.45	.887	.198	4.03	4.87	2	5
	b) Discussion-training (A128).	6	3.83	1.472	.601	2.29	5.38	2	5
	c) Web-based training (e-mail with lessons).	38	4.32	.962	.156	4.00	4.63	1	5
	d) None.	46	4.22	1.134	.167	3.88	4.55	1	5
	Total	110	4.27	1.048	.100	4.07	4.47	1	5
I don't use shredder for discarding the documents with sensitive information	a) Classroom-training (Eureka 2/3).	20	4.30	1.218	.272	3.73	4.87	1	5
	b) Discussion-training (A128).	6	3.83	1.472	.601	2.29	5.38	2	5
	c) Web-based training (e-mail with lessons).	38	4.39	1.028	.167	4.06	4.73	1	5
	d) None.	46	4.15	1.247	.184	3.78	4.52	1	5
	Total	110	4.25	1.175	.112	4.02	4.47	1	5
I save sensitive information in memory stick or external hard drive	a) Classroom-training (Eureka 2/3).	20	4.50	.946	.212	4.06	4.94	2	5
	b) Discussion-training (A128).	6	4.17	1.329	.543	2.77	5.56	2	5
	c) Web-based training (e-mail with lessons).	38	3.61	1.220	.198	3.20	4.01	1	5
	d) None.	46	3.43	1.276	.188	3.06	3.81	1	5
	Total	110	3.73	1.256	.120	3.49	3.96	1	5
I keep my desk clean from sensitive documents most of the time	a) Classroom-training (Eureka 2/3).	20	3.65	1.309	.293	3.04	4.26	1	5
	b) Discussion-training (A128).	6	4.17	1.169	.477	2.94	5.39	2	5
	c) Web-based training (e-mail with lessons).	38	3.89	1.060	.172	3.55	4.24	1	5
	d) None.	46	3.80	1.128	.166	3.47	4.14	1	5
	Total	110	3.83	1.132	.108	3.61	4.04	1	5
I write information about my work/research, or students in social networking sites (Facebook, twitter, MySpace)	a) Classroom-training (Eureka 2/3).	20	4.80	.696	.156	4.47	5.13	2	5
	b) Discussion-training (A128).	6	4.83	.408	.167	4.40	5.26	4	5
	c) Web-based training (e-mail with lessons).	38	4.84	.437	.071	4.70	4.99	3	5
	d) None.	46	4.83	.570	.084	4.66	5.00	2	5
	Total	110	4.83	.539	.051	4.73	4.93	2	5

Figure 17: Descriptives - questions group #2.

Test of Homogeneity of Variances

	Levene Statistic	df1	df2	Sig.
I don't lock the door of my office during my working hours, even if I am away	.787	3	106	.504
I open unexpected files or e-mail attachments or files, that I receive from unknown or known sender	.650	3	106	.585
I share sensitive information about my work with all my colleagues (such as information about projects, personal information about students, etc.)	.554	3	106	.646
I put my paper documents that contain sensitive information in the recycle bin for paper	2.430	3	106	.069
I don't use shredder for discarding the documents with sensitive information	1.194	3	106	.316
I save sensitive information in memory stick or external hard drive	2.740	3	106	.047
I keep my desk clean from sensitive documents most of the time	1.082	3	106	.360
I write information about my work/research, or students in social networking sites (Facebook, twitter, MySpace)	.169	3	106	.917

Figure 18: Test of Homogeneity of Variances - questions group #2.

		ANOVA				
		Sum of Squares	df	Mean Square	F	Sig.
I don't lock the door of my office during my working hours, even if I am away	Between Groups	5.247	3	1.749	1.372	.255
	Within Groups	135.125	106	1.275		
	Total	140.373	109			
I open unexpected files or e-mail attachments or files, that I receive from unknown or known sender	Between Groups	.441	3	.147	.275	.843
	Within Groups	56.731	106	.535		
	Total	57.173	109			
I share sensitive information about my work with all my colleagues (such as information about projects, personal information about students, etc.)	Between Groups	1.535	3	.512	.631	.596
	Within Groups	85.920	106	.811		
	Total	87.455	109			
I put my paper documents that contain sensitive information in the recycle bin for paper	Between Groups	1.998	3	.666	.599	.617
	Within Groups	117.820	106	1.112		
	Total	119.818	109			
I don't use shredder for discarding the documents with sensitive information	Between Groups	2.326	3	.775	.555	.646
	Within Groups	148.047	106	1.397		
	Total	150.373	109			
I save sensitive information in memory stick or external hard drive	Between Groups	17.602	3	5.867	4.033	.009
	Within Groups	154.217	106	1.455		
	Total	171.818	109			
I keep my desk clean from sensitive documents most of the time	Between Groups	1.517	3	.506	.388	.762
	Within Groups	138.201	106	1.304		
	Total	139.718	109			
I write information about my work/research, or students in social networking sites (Facebook, twitter, MySpace)	Between Groups	.024	3	.008	.026	.994
	Within Groups	31.695	106	.299		
	Total	31.718	109			

Figure 19: ANOVA table - questions group #2.

Multiple Comparisons

LSD

Dependent Variable	(I) Which training group where you in:	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval		
					Lower Bound	Upper Bound	
I don't lock the door of my office during my working hours, even if I am away	a) Classroom-training (Eureka 2/3).	b) Discussion-training (A128).	-.217	.526	.681	-1.26	.83
		c) Web-based training (e-mail with lessons).	.082	.312	.794	-.54	.70
		d) None.	.450	.302	.140	-.15	1.05
	b) Discussion-training (A128).	c) Web-based training (e-mail with lessons).	.298	.496	.549	-.69	1.28
	d) None.	.667	.490	.177	-.30	1.64	
	c) Web-based training (e-...	d) None.	.368	.248	.140	-.12	.86
I open unexpected files or e-mail attachments or files, that I receive from unknown or known sender	a) Classroom-training (Eureka 2/3).	b) Discussion-training (A128).	-.133	.341	.696	-.81	.54
		c) Web-based training (e-mail with lessons).	.121	.202	.550	-.28	.52
		d) None.	.048	.196	.808	-.34	.44
	b) Discussion-training (A128).	c) Web-based training (e-mail with lessons).	.254	.321	.430	-.38	.89
	d) None.	.181	.318	.570	-.45	.81	
	c) Web-based training (e-...	d) None.	-.073	.160	.649	-.39	.24
I share sensitive information about my work with all my colleagues (such as information about projects, personal information about students, etc.)	a) Classroom-training (Eureka 2/3).	b) Discussion-training (A128).	-.150	.419	.721	-.98	.68
		c) Web-based training (e-mail with lessons).	-.150	.249	.548	-.64	.34
		d) None.	.111	.241	.647	-.37	.59
	b) Discussion-training (A128).	c) Web-based training (e-mail with lessons).	.000	.396	1.000	-.78	.78
	d) None.	.261	.391	.506	-.51	1.04	
	c) Web-based training (e-...	d) None.	.261	.197	.189	-.13	.65
I put my paper documents that contain sensitive information in the recycle bin for paper	a) Classroom-training (Eureka 2/3).	b) Discussion-training (A128).	.617	.491	.212	-.36	1.59
		c) Web-based training (e-mail with lessons).	.134	.291	.646	-.44	.71
		d) None.	.233	.282	.412	-.33	.79
	b) Discussion-training (A128).	c) Web-based training (e-mail with lessons).	-.482	.463	.300	-1.40	.44
	d) None.	-.384	.458	.403	-1.29	.52	
	c) Web-based training (e-...	d) None.	.098	.231	.671	-.36	.56
I don't use shredder for discarding the documents with sensitive information	a) Classroom-training (Eureka 2/3).	b) Discussion-training (A128).	.467	.550	.398	-.62	1.56
		c) Web-based training (e-mail with lessons).	-.095	.326	.772	-.74	.55
		d) None.	.148	.317	.641	-.48	.78
	b) Discussion-training (A128).	c) Web-based training (e-mail with lessons).	-.561	.519	.282	-1.59	.47
	d) None.	-.319	.513	.536	-1.34	.70	
	c) Web-based training (e-...	d) None.	.243	.259	.351	-.27	.76
I save sensitive information in memory stick or external hard drive	a) Classroom-training (Eureka 2/3).	b) Discussion-training (A128).	.333	.561	.554	-.78	1.45
		c) Web-based training (e-mail with lessons).	.895	.333	.008	.23	1.56
		d) None.	1.065	.323	.001	.42	1.71
	b) Discussion-training (A128).	c) Web-based training (e-mail with lessons).	.561	.530	.292	-.49	1.61
	d) None.	.732	.524	.165	-.31	1.77	
	c) Web-based training (e-...	d) None.	.170	.264	.520	-.35	.69
I keep my desk clean from sensitive documents most of the time	a) Classroom-training (Eureka 2/3).	b) Discussion-training (A128).	-.517	.531	.333	-1.57	.54
		c) Web-based training (e-mail with lessons).	-.245	.315	.440	-.87	.38
		d) None.	-.154	.306	.615	-.76	.45
	b) Discussion-training (A128).	c) Web-based training (e-mail with lessons).	.272	.502	.589	-.72	1.27
	d) None.	.362	.496	.466	-.62	1.34	
	c) Web-based training (e-...	d) None.	.090	.250	.719	-.41	.59
I write information about my work/research, or students in social networking sites (Facebook, twitter, MySpace)	a) Classroom-training (Eureka 2/3).	b) Discussion-training (A128).	-.033	.255	.896	-.54	.47
		c) Web-based training (e-mail with lessons).	-.042	.151	.781	-.34	.26
		d) None.	-.026	.146	.859	-.32	.26
	b) Discussion-training (A128).	c) Web-based training (e-mail with lessons).	-.009	.240	.971	-.49	.47
	d) None.	.007	.237	.976	-.46	.48	
	c) Web-based training (e-...	d) None.	.016	.120	.894	-.22	.25

Figure 20: LSD - questions group #2.

Contrast Coefficients

Contrast	Which training group where you in:			
	a) Classroom- training (Eureka 2/3).	b) Discussion- training (A128).	c) Web-based training (e- mail with lessons).	d) None.
1	1	1	1	-3

Figure 21: Contrast coefficients for participants and non-participants.

In the same sequence, results for the second group of questions are presented. Test of homogeneity identifies one question, Q2_25, with significant difference between the levels. Overall ANOVA confirms same result, unlike LSD. Multiple comparisons identifies no significant difference between any levels (groups) for any question.

To verify obtained results, contrast test is done for both groups. Contrast test uses two-tailed significance value and proceeds two types of comparison: under equal-variance and unequal-variance estimates. First, we identify contrast for the levels (table 21). Here levels with positive coefficients will be compared with levels with negative ones. Sum of coefficients for each contrast should be equal to zero. So, levels, which identify participants of any training, will be grouped and compared with non-participants.

Contrast Tests

		Contrast	Value of Contrast	Std. Error	t	df	Sig. (2-tailed)
I write down my passwords in a piece of paper near my computer	Assume equal variances	1	.32	.776	.409	106	.684
	Does not assume equal variances	1	.32	.761	.417	33.866	.680
I save my passwords in my cell-phone or memory stick	Assume equal variances	1	.92	.604	1.520	106	.131
	Does not assume equal variances	1	.92	.496	1.852	65.978	.069
I use passwords that I can easily remember so I don't have to save them	Assume equal variances	1	-.66	.882	-.750	106	.455
	Does not assume equal variances	1	-.66	.933	-.709	25.501	.485
I don't have problem to tell my password to IT people if I am asked to	Assume equal variances	1	-.05	.681	-.075	106	.940
	Does not assume equal variances	1	-.05	.876	-.058	14.317	.954
I don't use Password-protected screen saver in my PC at work	Assume equal variances	1	1.00	1.043	.961	106	.339
	Does not assume equal variances	1	1.00	.964	1.039	43.216	.304
My password is shorter than 8 characters:	Assume equal variances	1	.25	.354	.695	106	.489
	Does not assume equal variances	1	.25	.350	.703	31.308	.487

Figure 22: Contrast test for training participants and non-participants - questions group #1.

For the first group of questions, contrast test (table 22) identifies question Q2_9 with significant value less than 0.1 (two-tailed value) for the unequal-variance estimate. Since significant value under the equal-variance estimates for the current question does not achieve significance, we cannot accept validity under unequal-variance estimates. Therefore, scores of question Q2_9 do not significantly differ between participants and non-participants.

For the second group, contrast test (table 23) finds two questions, whose scores are significantly different between participants and non-participants under both estimates.

Contrast Tests

		Contrast	Value of Contrast	Std. Error	t	df	Sig. (2-tailed)
I don't lock the door of my office during my working hours, even if I am away	Assume equal variances	1	1.49	.748	1.986	106	.050
	Does not assume equal variances	1	1.49	.662	2.243	71.977	.028
I open unexpected files or e-mail attachments or files, that I receive from unknown or known sender	Assume equal variances	1	.16	.485	.321	106	.748
	Does not assume equal variances	1	.16	.407	.383	69.793	.703
I share sensitive information about my work with all my colleagues (such as information about projects, personal information about students, etc.)	Assume equal variances	1	.63	.596	1.061	106	.291
	Does not assume equal variances	1	.63	.604	1.047	36.922	.302
I put my paper documents that contain sensitive information in the recycle bin for paper	Assume equal variances	1	-.05	.698	-.076	106	.940
	Does not assume equal variances	1	-.05	.822	-.065	16.578	.949
I don't use shredder for discarding the documents with sensitive information	Assume equal variances	1	.07	.783	.091	106	.927
	Does not assume equal variances	1	.07	.876	.082	20.691	.936
I save sensitive information in memory stick or external hard drive	Assume equal variances	1	1.97	.799	2.463	106	.015
	Does not assume equal variances	1	1.97	.835	2.357	24.611	.027
I keep my desk clean from sensitive documents most of the time	Assume equal variances	1	.30	.756	.395	106	.694
	Does not assume equal variances	1	.30	.769	.388	28.802	.701
I write information about my work/research, or students in social networking sites (Facebook, twitter, MySpace)	Assume equal variances	1	.00	.362	-.008	106	.994
	Does not assume equal variances	1	.00	.347	-.008	52.723	.994

Figure 23: Contrast test for training participants and non-participants - questions group #2.

Relying on results of our statistical analysis, we can conclude that there is no significant difference between awareness level of employees before and after the training program. Moreover, we did not identify any of training program, which scores significantly higher than others.

6 Future Work

While working with this thesis, we found multiple possibilities to extend our work beyond the available time for the master thesis. This chapter will summarize the future work.

The methodology of questionnaire design can be significantly improved by extending the concept. In the input framework, sampling of a population can be added. It will increase quality of the evaluation results. Moreover, variables specification can be implemented in the development stage of the questionnaire. In other words, methodology can be improved by connecting it with the following statistical analysis. For our future work, the main goal is to make questionnaire design and statistical analysis interconnected. It will work positively in both directions. Moreover, it will guarantee reliable results.

For the future work, we would also like to have a collaboration work with authors of Information Security Culture Decision Support System [43, 35, 36]. We think that our methodology of questionnaire can be integrated and used in this system. Certainly, it needs to be improved and adopted for current purposes of the system. Moreover, the possibility to have such collaboration in a future, was discussed on the meetings with the supervisor of this master thesis, Prof. Dr. Bernhard M. Hämmerli.

Furthermore, we discussed in the state-of-the-art that questionnaire can be also used to put some ideas into respondents' minds. So, for a future work, we want to investigate how we can use survey instead of the trainings. Survey will be, first of all, addressed to make an impact on respondent's awareness, not to measure it.

For the measurement of information security awareness is clear that new survey is needed. The approach, chosen for such measurement, is completely reliable and significant, in our opinion. But what has failed that it is a performance. Therefore, we recommend to do the same survey for a future work. For a new survey, training programs should be available and target group carefully selected. Survey should be well-developed and tested before going to the target group. Rate of the participation in the trainings should be managed and, if possible, increased. Data should be collected from the same target group before and after the training and then statistically analyzed. Also, comparison between training programs should be based not just on awareness level, but on programs' design as well. Therefore, programs should cover same issues and be designed on the same background. Just their representation can vary (like web-based, discussion trainings etc.).

7 Conclusion

In this master thesis we designed and tested new methodology of questionnaire design and performed re-evaluation of an available survey. In this chapter we summarize our findings and achievements. First, we discuss a new methodology of questionnaire design, then, we discuss evaluation results of survey, which measures effectiveness of information security awareness program.

We have designed a new methodology of questionnaire design, which unlike existing approaches, tests and verifies questionnaire on its development stage. The methodology is designed to be applicable for wide range of survey topics. It is simple to use and it does not require specific knowledge or tools. It is based on the same procedures as a survey: questionnaires and semi-structured interviews. The methodology includes a framework, which is basically guidelines, for first-version questionnaire and testing and verification part. Testing and verification part is addressed to improve a quality of the questionnaire and make it ready to be send out to a target group. Testing and verification part includes two loops, which can be used separately and together. Except verification, this part additionally allows to design response options to the questions. To improve performance of testing and verification, trusted groups are introduced. Methodology also includes suggestions according to social relationship between researcher and target group, introduction letter.

We have applied designed methodology for questionnaires from a survey, chosen for re-evaluation. Results include improved version of a questionnaire, which is made based on previous work, semi-structured interviews and data collected from trusted group, first loop of iterations. Moreover, analysis of questionnaire design, as it was predicted, has a huge impact on the statistical evaluation.

Statistical analysis of a survey shows that information security training did not significantly change the level of awareness among the employees of Gjøvik University College. Moreover, we did not identify a significant difference between performances of any of the programs. Nevertheless, we suggest to not rely on these results completely. According to our complex analysis, we propose to make the same survey again. But this time, more carefully design questionnaires and make a training not just on voluntary basis. Evaluation results are not considered by us as absolutely relevant, because of small samples for each group. Moreover, the data is completely out-to-date, two years is quite an age for a survey.

Bibliography

- [1] *IBM SPSS Statistics tutorial*.
- [2] Enalyzer survey solution. <http://www.enalyzer.com/>, 2013.
- [3] PSPP - a free replacement for the proprietary program spss. <http://www.gnu.org/software/pspp/>, 2013.
- [4] ANTONIUS, R. *Interpreting Quantitative Data with IBM SPSS Statistics*. SAGE Publications Limited, 2012.
- [5] BUCKINGHAM, A., AND SAUNDERS, P. *The survey methods workbook: From design to analysis*. Polity, 2004.
- [6] CHIPPERFIELD, J. O., AND STEEL, D. G. Efficiency of split questionnaire surveys. *Journal of Statistical Planning and Inference* 141, 5 (2011), 1925–1932.
- [7] DA VEIGA, A., AND ELOFF, J. A framework and assessment instrument for information security culture. *Computers & Security* 29, 2 (2010), 196–207.
- [8] DEL GRECO, L., WALOP, W., AND MCCARTHY, R. H. Questionnaire development: 2. validity and reliability. *CMAJ: Canadian Medical Association Journal* 136, 7 (1987), 699.
- [9] DILLMAN, D. A. *Mail and Internet surveys: The tailored design method–2007 Update with new Internet, visual, and mixed-mode guide*. Wiley, 2011.
- [10] DOLNICAR, S., AND GRÜN, B. “translating” between survey answer formats. *Journal of Business Research* (2012).
- [11] ERGU, D., AND KOU, G. Questionnaire design improvement and missing item scores estimation for rapid and efficient decision making. *Annals of Operations Research* 197, 1 (2012), 5–23.
- [12] FAN, W., AND YAN, Z. Factors affecting response rates of the web survey: A systematic review. *Computers in Human Behavior* 26, 2 (2010), 132–139.
- [13] FELDE, M. Change management applied to information security culture. Master’s thesis, Gjøvik University College, 2009.
- [14] FIELD, A., ET AL. *Discovering Statistics using IBM SPSS Statistics*. 2013.
- [15] GEORGE, D. *SPSS for Windows Step by Step: A Simple Study Guide and Reference, 18.0 Update, 10/e*. Pearson Education India, 2009.

- [16] GOOGLE. Create a Google form. <http://support.google.com/drive/bin/answer.py?hl=en&answer=87809>, 2013.
- [17] GORDON, L. A., AND LOEB, M. P. Budgeting process for information security expenditures. *Communications of the ACM* 49, 1 (2006), 121–125.
- [18] JAMES E. BARTLETT, II, J. W. K., AND HIGGINS, C. C. Organizational research: Determining appropriate sample size in survey research. *Information Technology, Learning, and Performance Journal* 19, 1 (2001).
- [19] JUNIPER, E., GUYATT, G., FERRIE, P., KING, D., ET AL. Development and validation of a questionnaire to measure asthma control. *European Respiratory Journal* 14, 4 (1999), 902–907.
- [20] KHALFAN, A. M. Information security considerations in IS/IT outsourcing projects: a descriptive case study of two sectors. *International Journal of Information Management* 24, 1 (2004), 29–42.
- [21] KIRKPATRICK, L. A., AND FEENEY, B. C. *A Simple Guide to IBM SPSS Statistics: For Versions 18.0*. Wadsworth Publishing Company, 2011.
- [22] KOTULIC, A. G., AND CLARK, J. G. Why there aren't more information security research studies. *Information & Management* 41, 5 (2004), 597–607.
- [23] KRUGER, H., AND KEARNEY, W. A prototype for assessing information security awareness. *Computers & Security* 25, 4 (2006), 289–296.
- [24] LACEY, D. *Managing the Human Factor in Information Security - How to win over staff and influence business managers*. John Wiley & Sons Ltd, 2009.
- [25] LEEDY, P. D., AND ORMROD, J. E. *Practical research, 10/e*. Prentice Hall Upper Saddle River, NJ, 2009.
- [26] MAKIENKO, I., AND BERNARD, E. K. Teaching applied value of marketing research: A questionnaire design project. *International Journal of Management Education* (2012).
- [27] MYERS, M. D., AND NEWMAN, M. The qualitative interview in IS research: Examining the craft. *Information and Organization* 17, 1 (2007), 2–26.
- [28] OPPENHEIM, A. N. *Questionnaire design, interviewing and attitude measurement*. Continuum, 2000.
- [29] PALLANT, J. *SPSS survival manual: A step by step guide to data analysis using SPSS*. Open University Press, 2010.
- [30] PRESTON, V. Questionnaire survey. In *International Encyclopedia of Human Geography*, E. in Chief: Rob Kitchin and N. Thrift, Eds. Elsevier, Oxford, 2009, pp. 46 – 52.

- [31] ROBERTS, M., STEWART, B. M., TINGLEY, D., LUCAS, C., LEDER-LUIS, J., ALBERTSON, B., GADARIAN, S., AND RAND, D. Topic models for open ended survey responses with applications to experiments. *Export BibTex Tagged XML topicmodelsopenendedexperiments. pdf* 1.
- [32] RÖMER, A., PACHE, M., WEISSHAHN, G., LINDEMANN, U., AND HACKER, W. Effort-saving product representations in design—results of a questionnaire survey. *Design Studies* 22, 6 (2001), 473–491.
- [33] SÁNCHEZ-FERNÁNDEZ, J., MUÑOZ-LEIVA, F., AND MONTORO-RÍOS, F. J. Improving retention rate and response quality in web-based surveys. *Computers in Human Behavior* 28, 2 (2012), 507–514.
- [34] SAURO, J., AND LEWIS, J. R. *Quantifying the user experience: Practical statistics for user research*. Morgan Kaufmann, 2012.
- [35] SCHLIENGER, T., AND TEUFEL, S. Information security culture - the socio-cultural dimension in information security management. *IFIP TC11 International Conference on Information Security (SEC 2002)* (2002).
- [36] SCHLIENGER, T., AND TEUFEL, S. Information security culture - from analysis to change. *Information Security South Africa - Proceedings of ISSA 2003*, 3 (2003), 183–195.
- [37] SCHLIENGER, T., AND TEUFEL, S. Tool supported management of information security culture: An application in a private bank. *The 20th IFIP International Information Security Conference (SEC 2005) - Security and Privacy in the Age of Ubiquitous Computing*, 20 (2005).
- [38] SMITS, N., AND VORST, H. Reducing the length of questionnaires through structurally incomplete designs: An illustration. *Learning and Individual Differences* 17, 1 (2007), 25–34.
- [39] SYSTEMS, C. C. I., AND SECURITY, D. Definition of sensitive information. http://www.colby.edu/administration_cs/its/policies/upload/Definition-of-Sensitive-Information.pdf, 2013.
- [40] SYSTEMS, C. R. Survey design. <http://www.surveysystem.com/sdesign.htm>, 2013.
- [41] VEHOVAR, V., LOZAR MANFREDA, K., KOREN, G., AND HLEBEC, V. Measuring ego-centered social networks on the web: Questionnaire design issues. *Social networks* 30, 3 (2008), 213–222.
- [42] VERMA, J. P. *Data Analysis in Management with SPSS Software*. Springer, 2013.
- [43] VESELI, I. Measuring the effectiveness of information security awareness program. Master's thesis, Gjøvik University College, Norway, 2011.
- [44] WAGNER III, W. E. *Using IBM® SPSS® Statistics for Research Methods and Social Science Statistics*. SAGE Publications, Incorporated, 2012.

- [45] WHITMAN, M. E., AND MATTORD, H. J. *Management of information security*. Course Technology Ptr, 2010.
- [46] ZIKMUND, W. G., AND BABIN, B. J. *Exploring marketing research, 4/e*. South-Western Pub, 2007.

A Improvement of Questionnaire Design Using Semi-Structured Interviews

This section shows how actually questions in the questionnaire were changed after each conducted interview. In table5 shows how questions from the second questionnaire (see table1, column "Second Questionnaire") were transferred into the questions for the semi-structured interview. Next table6 shows how questions after first interview were transformed to the second and so on. In following tables, column 2, Q#, refers to the original question in the second questionnaire. The lack of number indicates that this question was added and does not have an analog in the original questionnaire. Also, question #30 was added to the semi-structured interview from the first original questionnaire. In our opinion, it is reasonable to include in the interview because of its content.

It should be taken into the account, that these interviews are addressed to demonstrate how the methodology of questionnaire design - updating versions (see fig.4), works. The final version presented here, cannot be used as a new survey.

ADD DEFINITION SENSITIVE INFORMATION Any information for which loss, alteration, misuse or disclosure could adversely affect the interests of the College or its administration, faculty, staff, students, applicants or relations therein

Full version of the abbreviations used in tables:

- PW - password
- IS - information security

#	Q#	Interview Questions	Comments
1	Q3	In which department do you work?	
2	Q4	Are you a full-time or part-time employee?	
3		Have you participate in IS training in 2011 at Gjøvik University College ?	If NO, skip 2 next
4	Q5	Do you remember which kind of training program were you in?	
5	Q6	Do you remember the reason why you couldn't participate in such training?	
6	Q7	Do you think such IS training is important?	
7	Q8-9	Do you keep your PW somewhere in a written form?	If NO, skip next
8	Q8-9	In which written form? In cellphone, sticker, paper, on memory stick etc.?	
9	Q10	Do you use easy to remember PWs? So you don't need to store them anywhere?	

10	Q11	Do feel free to tell your PW to the IT department if they ask you to?	
11	Q12	For your PWs, which methods do you choose to create them?	
12	Q12	What do you think, will it be a good PW if you take for example line from a song and then first initials?	
13	Q13-14	Do you use different PW for private and working purposes?	
14	Q13-14	Do you use same PW to different accounts?	
15	Q15	Do you have PW-protected screen saver at work?	
16	Q15	Do you use it?	
17	Q16	Do you lock the door to your office while you are leaving even for a short time?	
18	Q17	When you receive unexpected emails from unknown senders, do you open such emails? Do you also download unknown attachments?	
19	Q18	Do you share information such as projects, personal information about students with you colleagues?	
20	Q19	Do you have any conversations about how to protect such information with your co-workers?	
21	Q20	What do you think IS policies might delay or disturb your regularly work?	
22	Q21	Who is in your opinion responsible for the IS at Gjøvik University College ?	
23	Q22	If you see some of your colleagues breaching IS rules, what are your actions?	
24	Q23-24	To utilize paper documents with sensitive information, what do you do?	
25	Q25	Do you store on external devices sensitive information you work with?	
26	Q26	When you need to work with some documents containing sensitive information, do you keep it after work on your desk?	
27	Q27	Do you share information about your work in social networks?	
28	Q28	Do you actually think more about IS after that training?	if answer on question#3 is YES
29	Q29	How many characters contain your PW?	
30		What is in your opinion a good PW example?	

Table 5: Questions for the first semi-structured interview.

#	Q#	Interview Questions	Comments
1	Q3	In which department do you work?	
2	Q4	Are you a full-time or part-time employee?	
3		Have you participate in IS training in 2011 at Gjøvik University College ?	If NO, skip 2 next
4	Q5	Do you remember which kind of training program were you in?	

5	Q6	Do you remember the reason why you couldn't participate in such training?	
6	Q7	Do you think such IS training is important?	
7	Q28	Do you actually think more about IS after that training?	If answer on question#3 is YES
8	Q10	Do you use easy to remember PW?	
9	Q8-9	Do you keep your PW somewhere in a written form?	If YES,next
10	Q8-9	In which written from? In cellphone, on sticker, memory stick etc.?	
11	Q8-9	Which kind of PW do you keep in a written from? F.e., which changes often, long and complicated etc.?	
12	Q13	Do you use different PWs for private and working purposes?	
13	Q13	Do you use same PW for different accounts?	
14	Q12	For your PWs, which methods do you choose to create it?	
15	Q12	What do you think, it will be a good PW if you take, for example, line from a song and then keep first initials?	
16	Q15	Do you have a PW-protected screen saver at work?	
17	Q15	Do you use it?	
18	Q29	How many characters does your usual PW have?	
19		What is in your opinion a good example of a PW?	
20	Q11	Do you feel free to tell your PW to the IT department if they ask you to?	
21	Q16	Do you lock the door to your office while you are leaving even for a short time?	
22	Q17	When you receive unexpected emails from unknown senders, do you open such emails? Do you also download unknown attachments?	
23	Q27	Do you share information related to your work in some social networks?	
24	Q25	Do you work at home sometimes?	If YES, next
25	Q25	Do you bring at home some sensitive information from job to work on?	If YES, next
26	Q25	Do you bring it in printed from or on external digital devices?	
27	Q26	When you need to work with documents which contain sensitive information, where do you keep them after you finished the work?	
28	Q23-24	To utilize paper documents with sensitive information, what do you do?	
29	Q20	What do you think, IS policies delay or disturb your regularly work?	
30	Q19	Do you have any conversations of how to protect sensitive information with your co-workers?	
31	Q22	Have you ever seen some of your colleagues breaching IS rules?	If YES, next
32	Q22	What were your actions?	
33	Q21	Who do you think is responsible for the IS at Gjøvik University College ?	

Table 6: Questions for the second semi-structured interview.

#	Q#	Interview Questions	Comments
1	Q3	In which department do you work?	
2	Q4	Are you a full-time or part-time employee?	
3		Have you participate in IS training in 2011 at Gjøvik University College ?	If NO, skip 2 next
4	Q5	Do you remember which kind of training program were you in?	
5	Q6	Do you remember the reason why you couldn't participate in such training?	
6	Q7	Do you think such IS training is important?	
7	Q28	Do you actually think more about IS after that training?	If answer on question#3 is YES
8	Q10	Do you use easy to remember PW?	
9	Q8-9	Do you keep your PW somewhere in a written form?	If YES, next
10	Q8-9	In which written from? In cellphone, on sticker, memory stick etc.?	
11	Q8-9	Which kind of PW do you keep in a written from? F.e., which changes often, long and complicated etc.?	
12	Q13	Do you use different PWs for private and working purposes?	
13	Q13	Do you use same PW for different accounts?	
14	Q12	For your PWs, which methods do you choose to create it?	
15	Q12	What do you think, it will be a good PW if you take, for example, line from a song and then keep first initials?	
16	Q15	Do you have a PW-protected screen saver at work?	
17	Q15	Do you use it?	
18	Q29	How many characters does your usual PW have?	
19		What is in your opinion a good example of a PW?	
20	Q11	Do you feel free to tell your PW to the IT department if they ask you to?	
21	Q27	Do you share information related to your work in some social networks?NOT Fronter	
22	Q16	Do you lock the door to your office while you are leaving even for a short time?	
23	Q17	When you receive unexpected emails from unknown senders, do you open such emails? Do you also download unknown attachments?	
24	Q25	Do you work at home sometimes?	If YES, next
25	Q25	Do you bring at home some sensitive information from job to work on?	If YES, next
26	Q25	Do you bring it in printed from or on external digital devices?	
27	Q26	When you need to work with documents which contain sensitive information, where do you keep them after you finished the work?	

28	Q23-24	To utilize paper documents with sensitive information, what do you do?	
29	Q23-24	Do you follow the rules stated in IS policy of how to utilize such documents?	
30	Q20	What do you think, IS policies delay or disturb your regularly work?	
31	Q19	Do you have any conversations of how to protect sensitive information with your co-workers?	
32	Q22	Have you ever seen some of your colleagues breaching IS rules?	If YES, next
33	Q22	What were your actions?	
34	Q21	Who do you think is responsible for the IS at Gjøvik University College ?	

Table 7: Questions for the third semi-structured interview.

#	Q#	Interview Questions	Comments
1	Q3	In which department do you work?	
2	Q4	Are you a full-time or part-time employee?	
3		Have you participate in IS training in 2011 at Gjøvik University College ?	If NO, skip 2 next
4	Q5	Do you remember which kind of training program were you in?	
5	Q6	Do you remember the reason why you couldn't participate in such training?	
6	Q7	Do you think such IS training is important?	
7	Q28	Do you actually think more about IS after that training?	If answer on question#3 is YES
8	Q10	Do you use easy to remember PW?	
9	Q8-9	Do you keep your PW somewhere in a written form?	If YES, next
10	Q8-9	In which written from? In cellphone, on sticker, memory stick etc.?	
11	Q8-9	Why do you keep it in a written form?	
12	Q13	Do you use same PW for different accounts?	
13	Q12	How do you usually create a new PW?	
14	Q12	What do you think, it will be a good approach to create a PW?	
15	Q15	Do you have a PW-protected screen saver at work?	
16	Q15	Do you use it?	
17	Q29	How many characters does your usual PW have?	
18		What is in your opinion a good example of a PW?	
19	Q11	Do you feel free to tell your PW to the IT department if they ask you to?	
20	Q27	Do you share information related to your work in some social networks?NOT Fronter	
21	Q25	Do you work at home sometimes?	If YES, next
22	Q25	Do you bring at home some sensitive information from job to work on?	If YES, next

23	Q25	Do you bring it in printed form or on external digital devices?	
24	Q26	When you need to work with documents which contain sensitive information, where do you keep them after you finished the work?	
25	Q16	Do you lock the door to your office while you are leaving even for a short time?	
26	Q17	When you receive unexpected emails from unknown senders, do you open such emails? Do you also download unknown attachments?	
27	Q23-24	To utilize paper documents with sensitive information, what do you do?	
28	Q23-24	Do you follow the rules stated in IS policy of how to utilize such documents?	
29	Q20	What do you think, IS policies delay or disturb your regularly work?	
30	Q19	Do you have any conversations of how to protect sensitive information with your co-workers?	
31	Q22	Have you ever seen some of your colleagues breaching IS rules?	If YES, next
32	Q22	What were your actions?	
33	Q21	Who do you think is responsible for the IS at Gjøvik University College ?	

Table 8: Questions for the fourth semi-structured interview.

#	Q#	Interview Questions	Comments
1	Q3	In which department do you work?	
2	Q4	Are you a full-time or part-time employee?	
3		Have you participate as a listener in IS training in 2011 at Gjøvik University College ?	If NO, skip 2 next
4	Q5	Do you remember which kind of training program were you in?	
5	Q6	Do you remember the reason why you couldn't participate in such training?	
6	Q28	Do you actually think more about IS after that training?	If answer on question#3 is YES
7	Q7	Do you think such IS training is important?	
8	Q10	Do you use easy to remember PW?	
9	Q8-9	Do you keep your PW somewhere in a written form?	If YES, 3 next
10	Q8-9	In which written form? In cellphone, on sticker, memory stick etc.?	
11	Q8-9	What is the reason for you to keep your PW somewhere?	
12	Q13	Do you use same PW for different kind of accounts?	
13	Q12	How do you usually create a new PW?	
14	Q15	Do you have a PW-protected screen saver at work? Login screen on you PC where you have to type your user name and PW	

15	Q15	Do you use it each time your PC was left without usage for a while?	
16	Q29	Approximately, how many characters does your usual PW have?	
17		What is in your opinion a good example of a PW?	
18	Q12	What do you think, it will be a good approach to create a PW?	
19	Q11	Do you feel free to tell your PW to the IT department if they ask you to?	
20	Q27	Do you share information related to your work in some social networks?NOT Fronter	
21	Q25	Do you work at home sometimes?	If YES, next
22	Q25	Do you bring at home some sensitive information from job to work on?	If YES, next
23	Q25	In which form do you take it home? (in electronic, papers, access external server from home through VPN)	
24	Q26	When you need to work with sensitive documents in your office, where do you keep them after you finished the work?	
25	Q23-24	To utilize paper documents with sensitive information, what do you do?	
26	Q23-24	Are you aware of special rules stated in IS policy about which documents and how should be utilized?	
27	Q16	Do you lock the door in your office while you are leaving for a short time?	
28	Q17	When you receive unexpected emails, do you open such emails?	
29	Q17	Do you also download unknown attachments?	
30	Q20	What do you think, IS policies delay or disturb your regularly work?	
31	Q19	Do you discuss with the colleagues how to protect sensitive information you work with?	
32	Q22	Have you ever seen some of your colleagues breaching IS rules?	If YES, next
33	Q22	What were your actions?	
34	Q21	Who do you think is responsible for the IS at Gjøvik University College ?	

Table 9: Questions for the fifth semi-structured interview.

#	Q#	Interview Questions	Comments
1	Q3	In which department do you work?	
2	Q4	Are you a full-time or part-time employee?	
3		Have you participate as a listener in IS training in 2011 at Gjøvik University College ?	If NO, skip 2 next
4	Q5	Do you remember which kind of training program were you in?	
5	Q6	Do you remember the reason why you couldn't participate in such training?	
6	Q28	Do you actually think more about IS after that training?	If answer on question#3 is YES
7	Q7	Do you think IS training is important?	
8	Q10	Do you use easy to remember PW?	

9	Q8-9	Do you keep your PW somewhere in a written form?	If YES, 3 next
10	Q8-9	In which written form? In cellphone, on sticker, memory stick etc.?	
11	Q8-9	What is the reason for you to keep your PW somewhere?	
12	Q13	Do you use same PW for different kind of accounts?	
13	Q29	Approximately, how many characters does your usual working PW have?	
14		What is in your opinion a good example of a PW?	
15	Q12	How do you usually create a new PW?	
16	Q12	What do you think, it will be a good approach to create a PW?	
17	Q11	Do you feel free to tell your PW to the IT department if they ask you to?	
18	Q15	Does your PC go automatically to login page after a while without usage?	
19	Q15	Do you lock your PC each time you are not leaving working place?	
20	Q16	Do you lock the door into your office while you are leaving even for a short time?	
21	Q27	Do you share information related to your work in some social networks? NOT Fronter	
22	Q17	When you receive unexpected emails, do you open such emails?	
23	Q17	Do you also download unknown attachments?	
24	Q25	Do you work at home sometimes?	If YES, next
25	Q25	Do you bring at home some sensitive information from office?	If YES, next
26	Q25	In which form do you bring it home? (in electronic, papers, access external server from home through VPN)	
27	Q26	When you need to work with sensitive documents in your office, where do you keep them after you finished the work?	
28	Q23-24	To destroy the paper documents with sensitive information, what do you do?	
29	Q23-24	Are you aware of special rules stated in IS policy about which documents and how should be utilized?	
30	Q20	What do you think, IS policies delay or disturb your regularly work?	
31	Q19	Do you discuss with the colleagues how to protect sensitive information you work with?	
32	Q22	Have you ever seen some of your colleagues breaching IS rules?	If YES, next
33	Q22	What were your actions?	
34	Q21	Who do you think is responsible for the IS at Gjøvik University College ?	

Table 10: Questions for the sixth semi-structured interview.

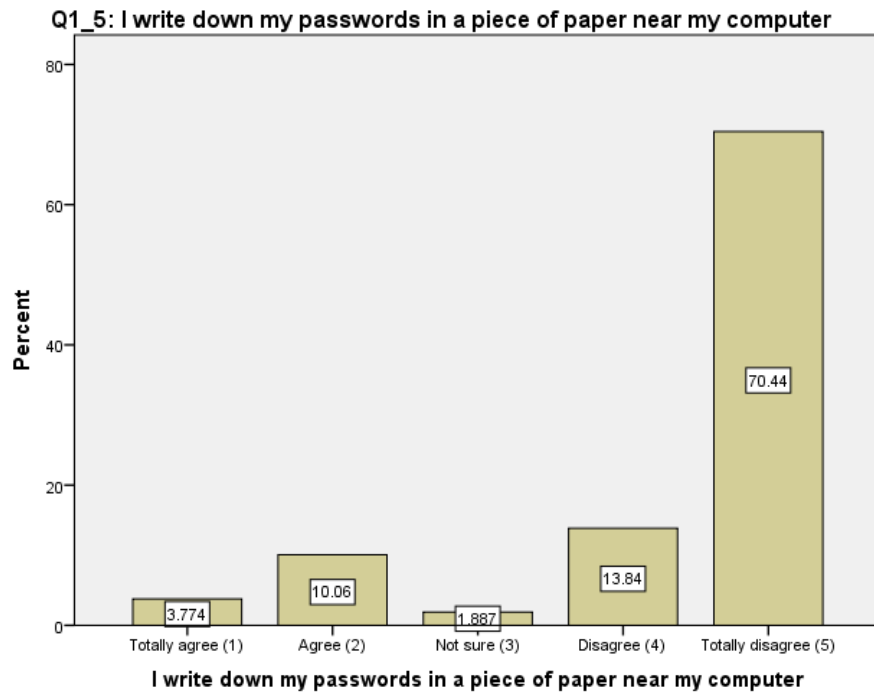
#	Q#	Interview Questions	Comments
1	Q3	In which department do you work?	

2	Q4	Are you a full-time or part-time employee?	
3		Have you participate as a listener in IS training in 2011 at Gjøvik University College ?	If NO, skip 3 next
4	Q5	Do you remember which kind of training program you were in?	
5	Q6	Do you remember the reason why you couldn't participate in this training?	
6	Q28	Do you actually think more about IS after you attended that training?	
7	Q7	Do you think IS training is important for the employees?	
8	Q10	Do you use easy to remember PW at work?	
9	Q8-9	Do you keep your PW somewhere in a written form?	If YES, 2 next
10	Q8-9	In which written from? In cellphone, on sticker, memory stick etc.?	
11	Q8-9	What is the reason for you to keep your PW somewhere?	
12	Q13	Do you use same PW for work and private accounts?	
13	Q29	Approximately, how many characters does your usual working PW have (if there is no limits for the characters number)?	
14		What is in your opinion a good example of a PW?	
15	Q12	Please, describe generally the technique you usually use to create a new PW?	
16	Q12	What do you think, it will be a good approach to create a PW?	
17	Q21	Who do you think is responsible for the IS at Gjøvik University College ?	
18	Q11	Do you feel free to tell your PW to the IT department if they ask you to?	
19	Q15	Does your PC go automatically to login page if you did not use it for a while?	
20	Q15	Do you lock your PC each time you are leaving working place?	
21	Q16	Do you lock the door into your office while you are leaving even for a short time?	
22	Q27	Do you share information related to your work in some social networks?NOT Fronter	
23	Q17	When you receive unexpected emails, do you open such emails?	
24	Q17	Do you also download unknown attachments?	
25	Q25	Do you work at home sometimes?	If YES, next
26	Q25	Do you bring at home some sensitive information from work?	If YES, next
27	Q25	In which form do you bring it home? (on external devices, printed papers or through the access to external server)	
28	Q25	If you use external server, do you connect to it through the VPN?	
29	Q26	When you need to work with sensitive documents in your office, where do you keep them after you finished the work?	
30	Q23-24	To destroy the paper documents with sensitive information, what do you do?	
31	Q23-24	Are you aware of special rules stated in IS policy about which documents and how should be utilized?	

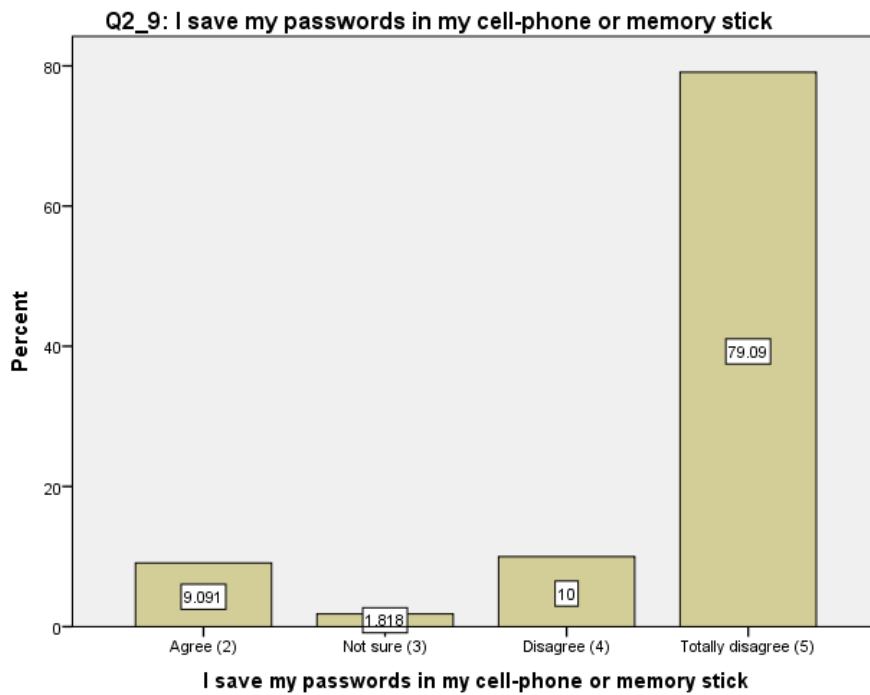
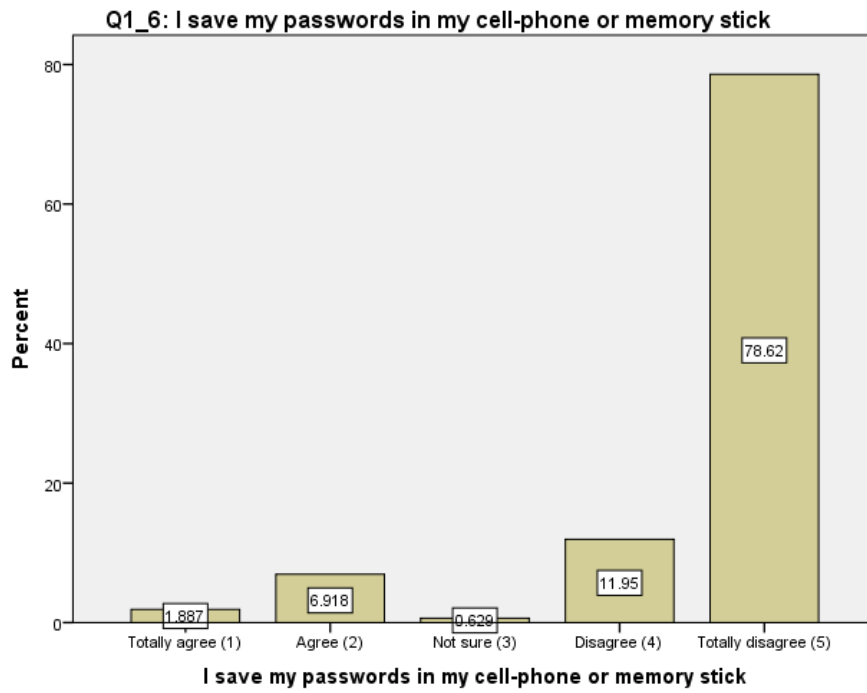
32	Q20	What do you think, IS policies delay or disturb your regularly work?	
33	Q19	Do you discuss with the colleagues how to protect sensitive information in some casual conversations?	
34	Q22	Have you ever seen some of your colleagues breaching IS rules?	If YES, next
35	Q22	What were your actions?	

Table 11: Questions for the seventh semi-structured interview.

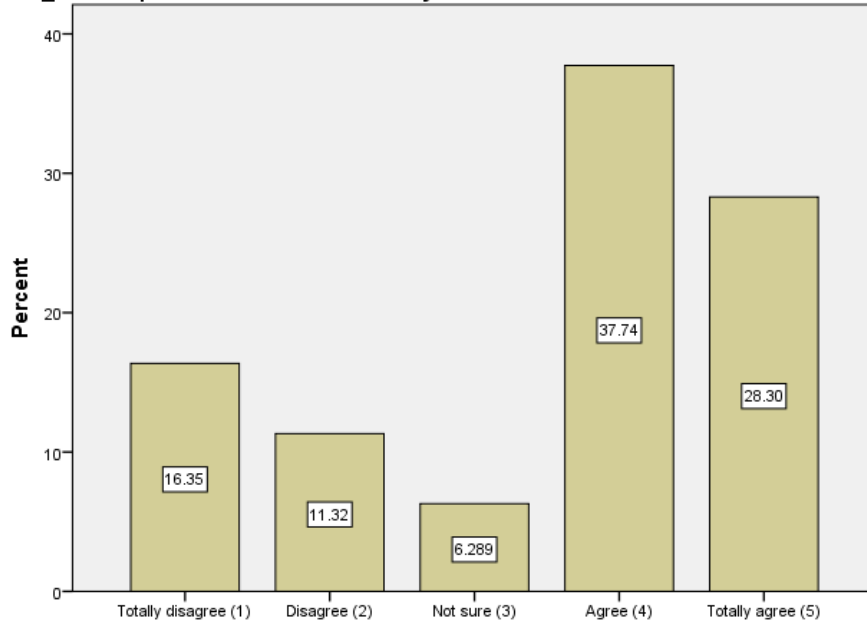
B Frequencies Distributions among Both Surveys





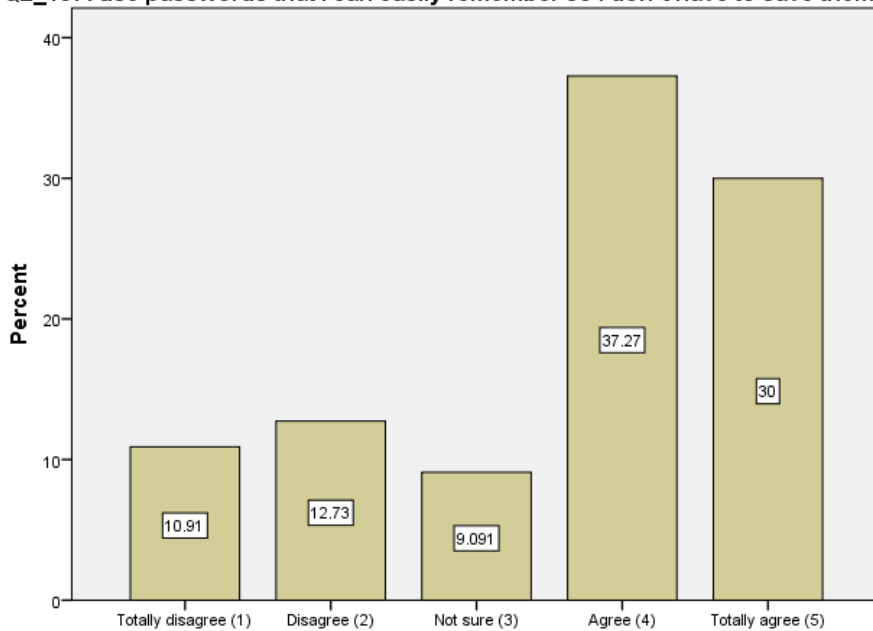


Q1_7: I use passwords that I can easily remember so I don't have to save them

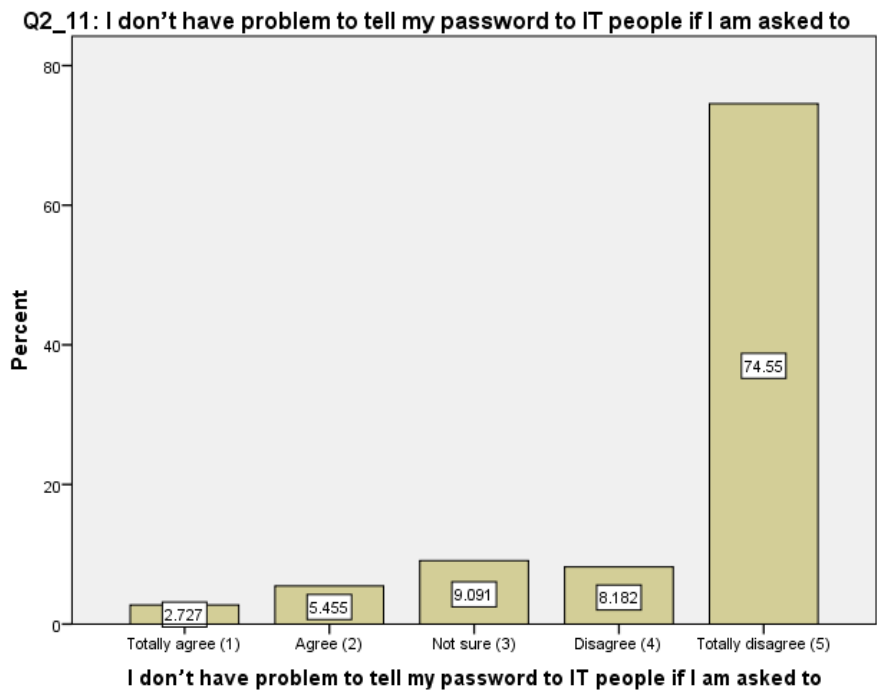
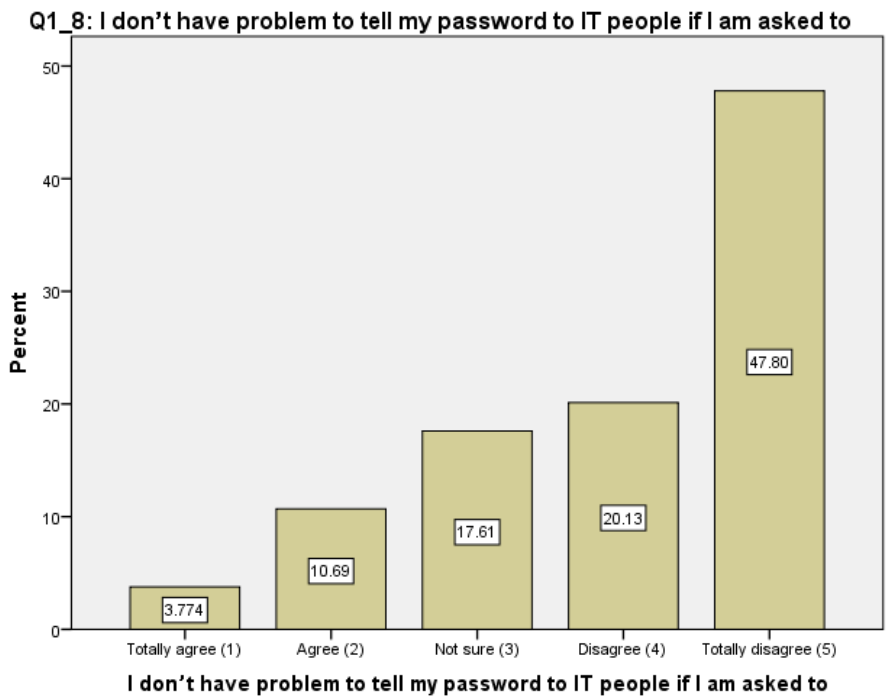


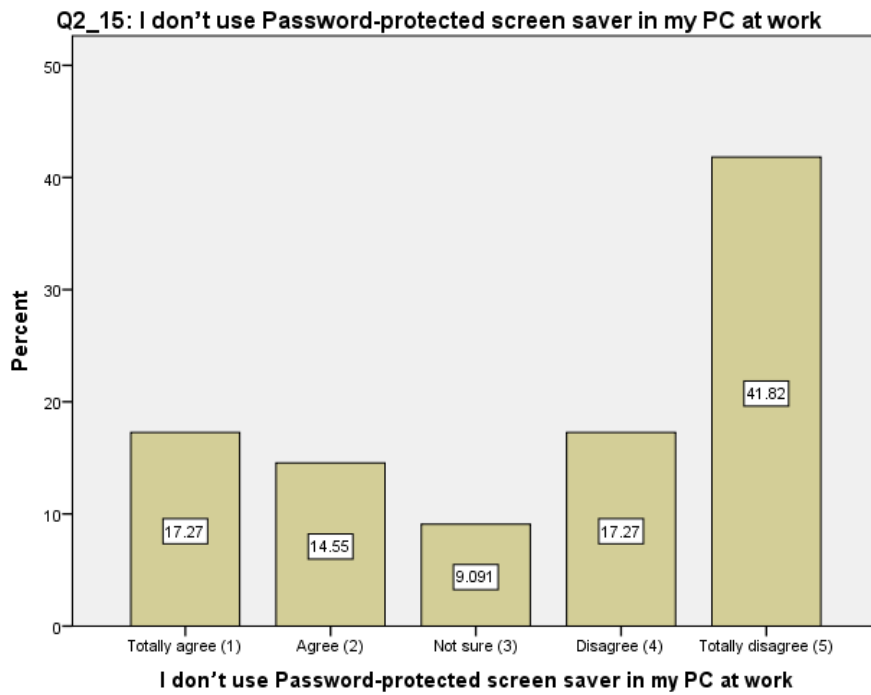
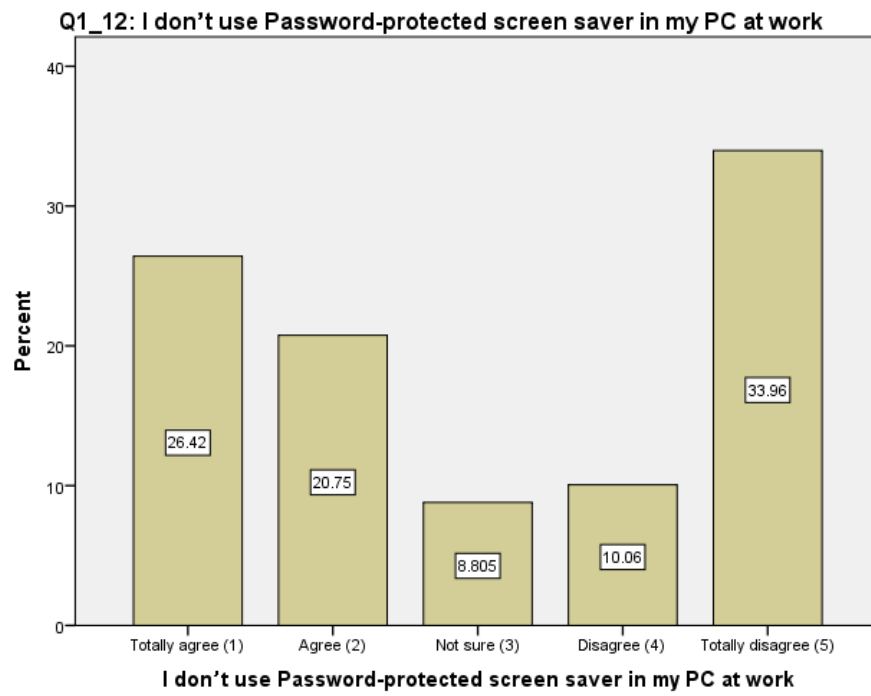
I use passwords that I can easily remember so I don't have to save them

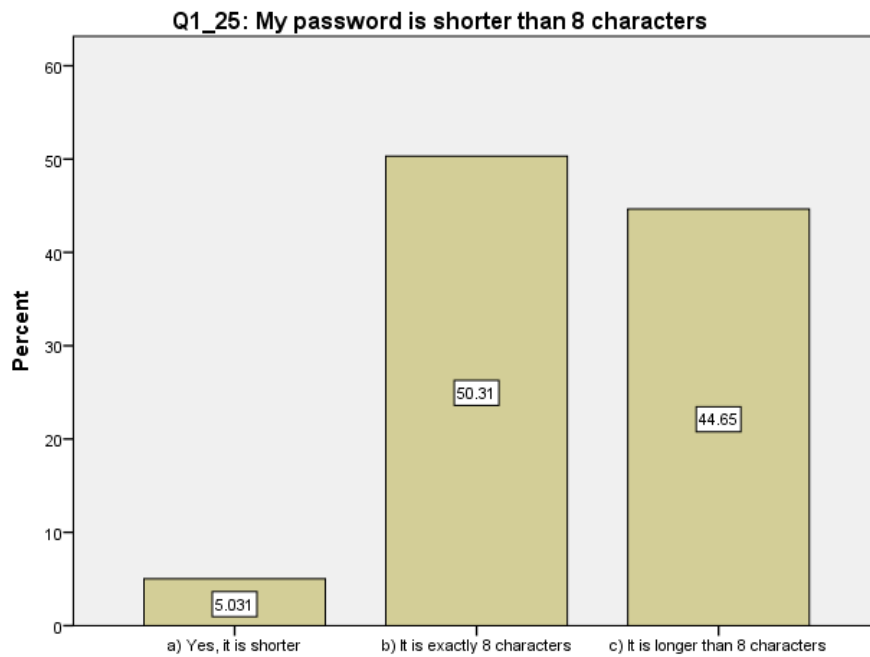
Q2_10: I use passwords that I can easily remember so I don't have to save them



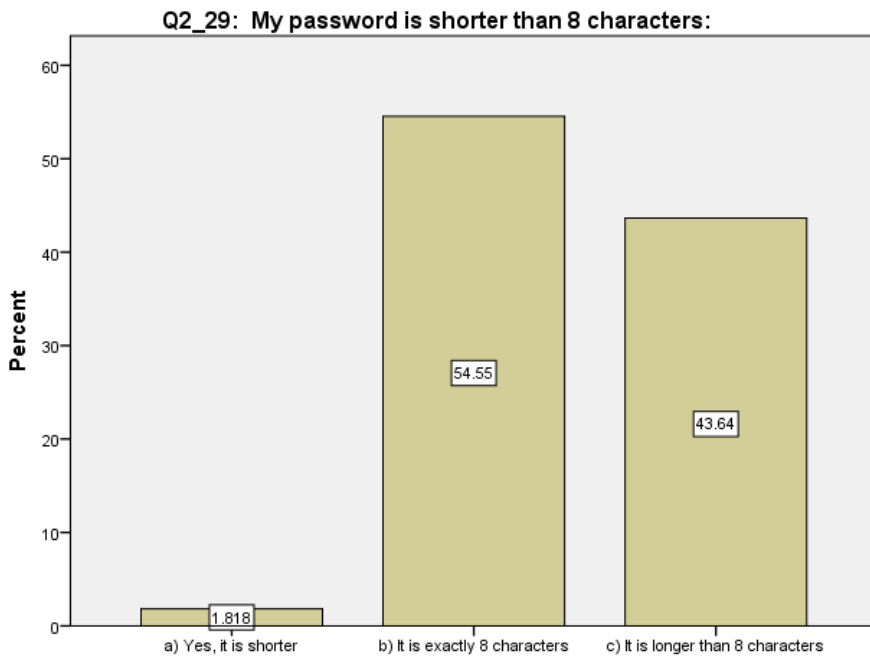
I use passwords that I can easily remember so I don't have to save them







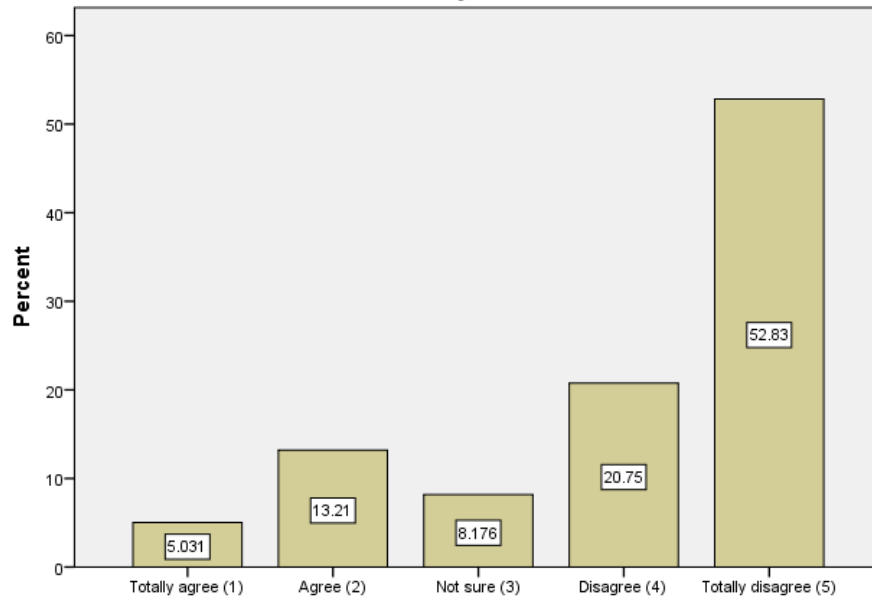
My password is shorter than 8 characters



My password is shorter than 8 characters:

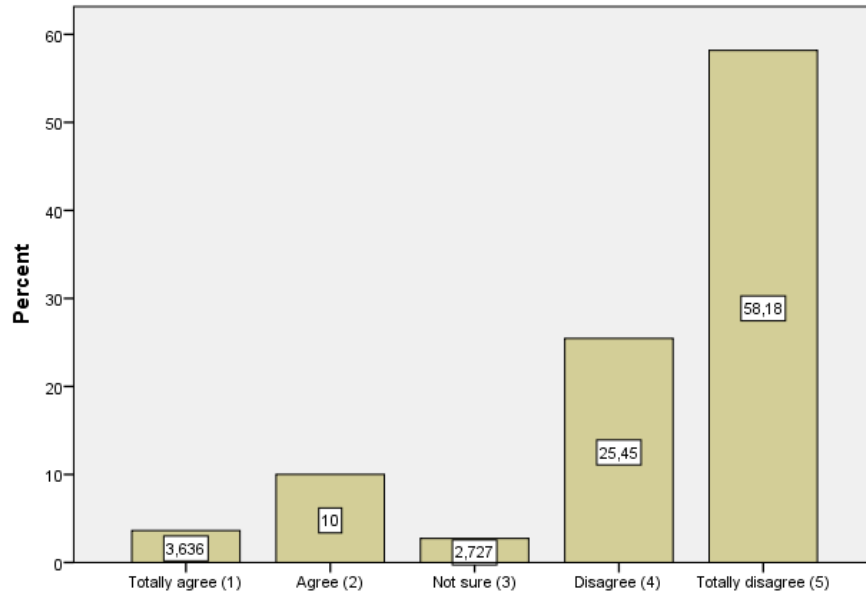
Figure 24: First group - Password Management.

Q1_13: I don't lock the door of my office during my working hours, even if I am away



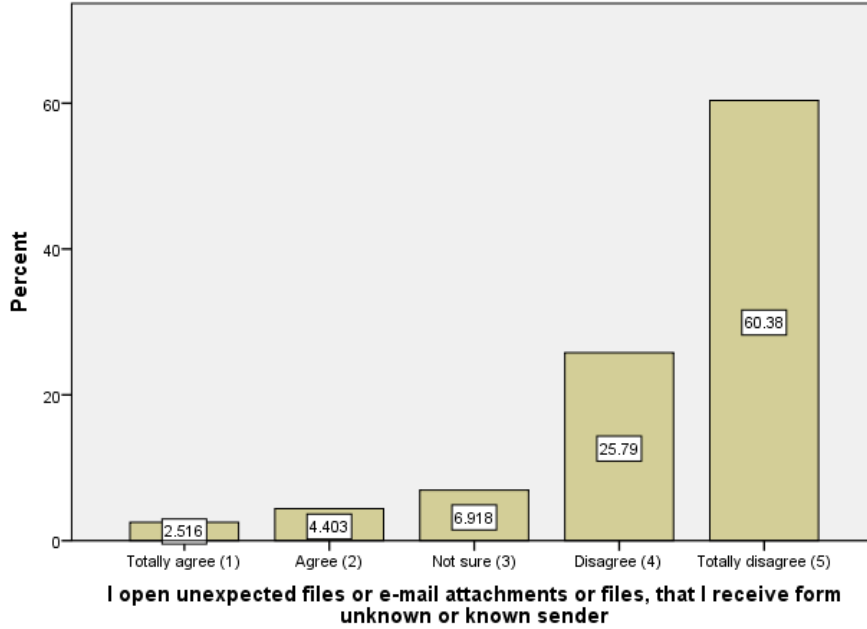
I don't lock the door of my office during my working hours, even if I am away

Q2_16: I don't lock the door of my office during my working hours, even if I am away

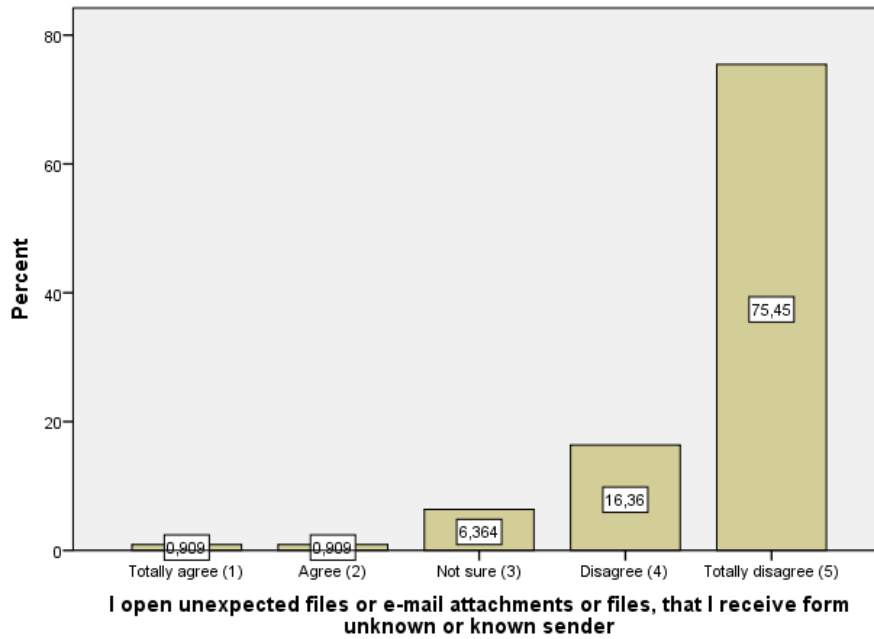


I don't lock the door of my office during my working hours, even if I am away

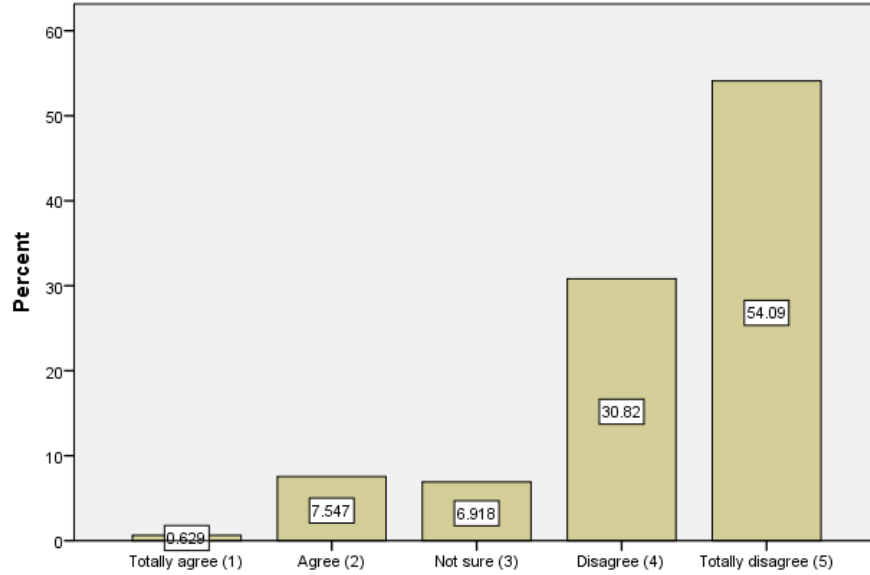
Q1_14: I open unexpected files or e-mail attachments or files, that I receive form unknown or known sender



Q2_17: I open unexpected files or e-mail attachments or files, that I receive form unknown or known sender

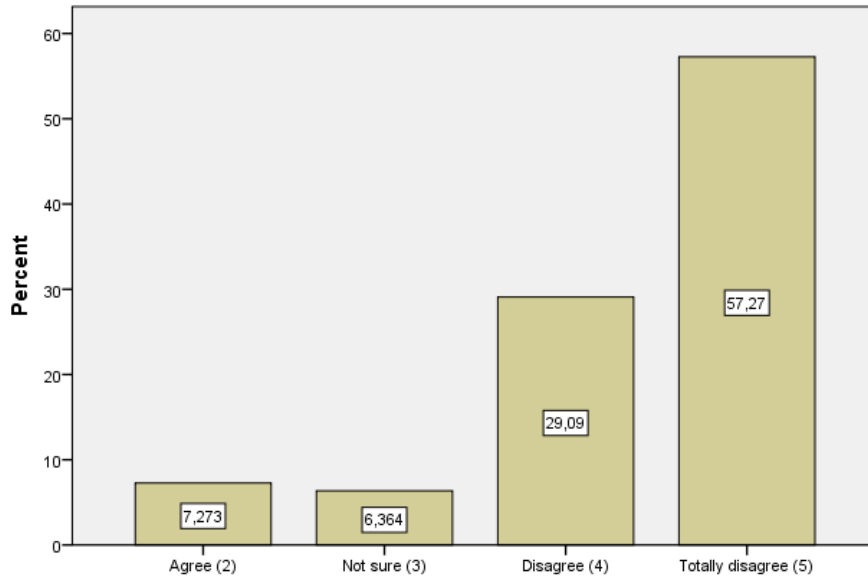


Q1_15: I share sensitive information about my work with all my colleagues (such as information about projects, personal information about students, etc.)



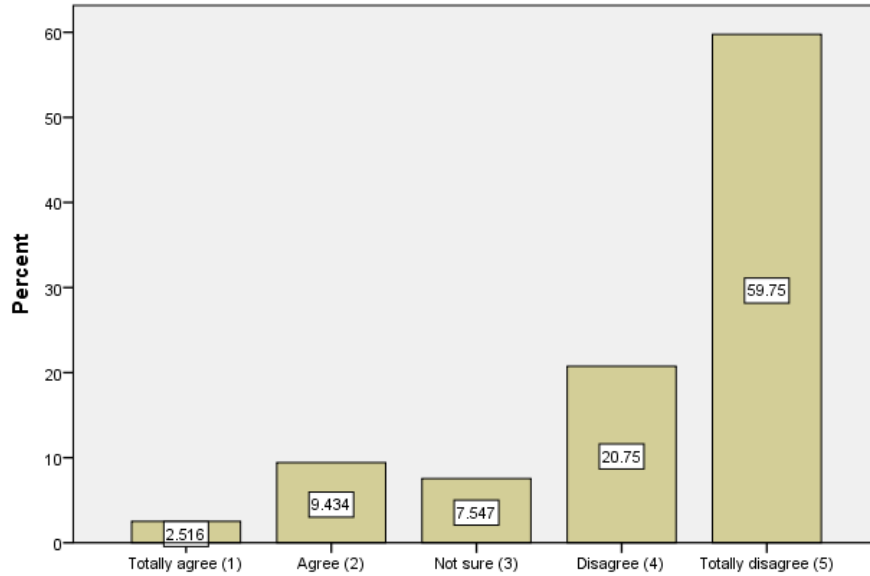
I share sensitive information about my work with all my colleagues (such as information about projects, personal information about students, etc.)

Q2_18: I share sensitive information about my work with all my colleagues (such as information about projects, personal information about students, etc.)



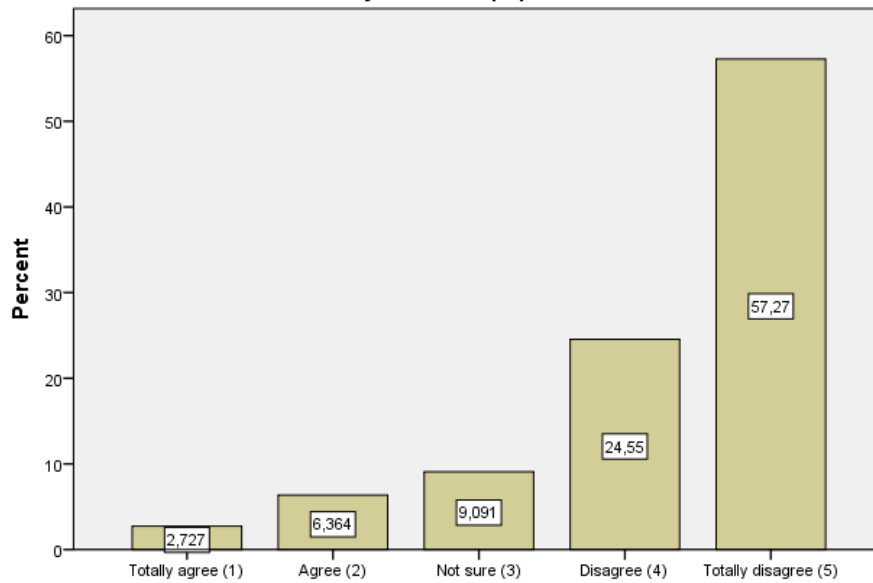
I share sensitive information about my work with all my colleagues (such as information about projects, personal information about students, etc.)

Q1_20: I put my paper documents that contain sensitive information in the recycle bin for paper



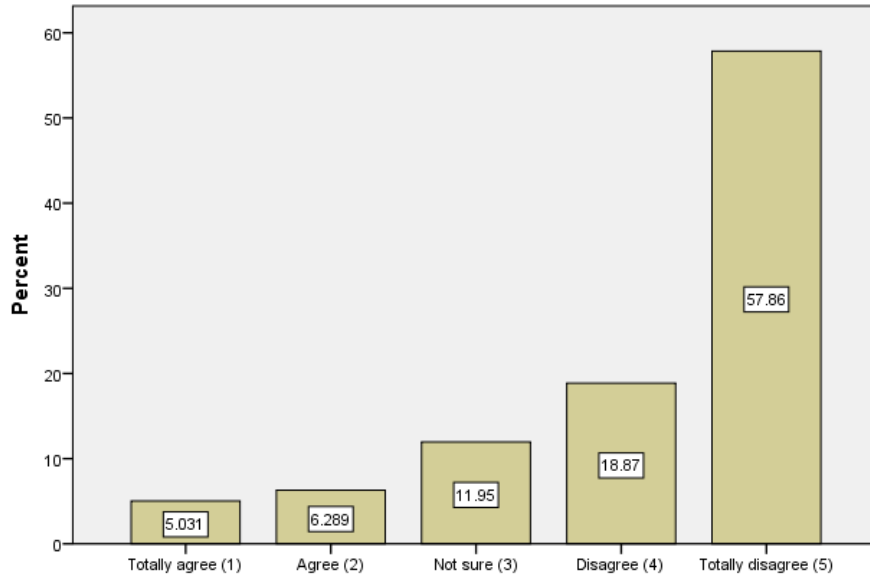
I put my paper documents that contain sensitive information in the recycle bin for paper

Q2_23: I put my paper documents that contain sensitive information in the recycle bin for paper



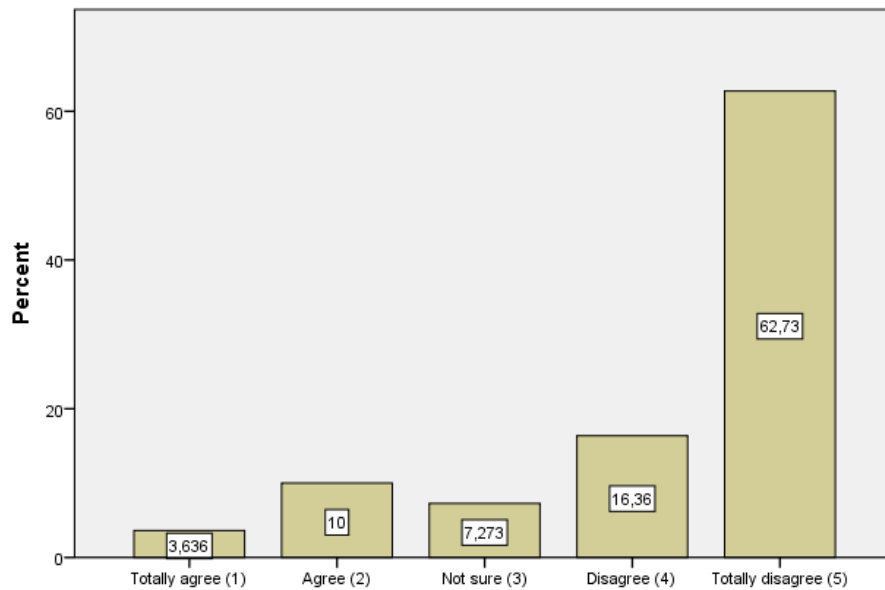
I put my paper documents that contain sensitive information in the recycle bin for paper

Q1_21: I don't use shredder for discarding the documents with sensitive information



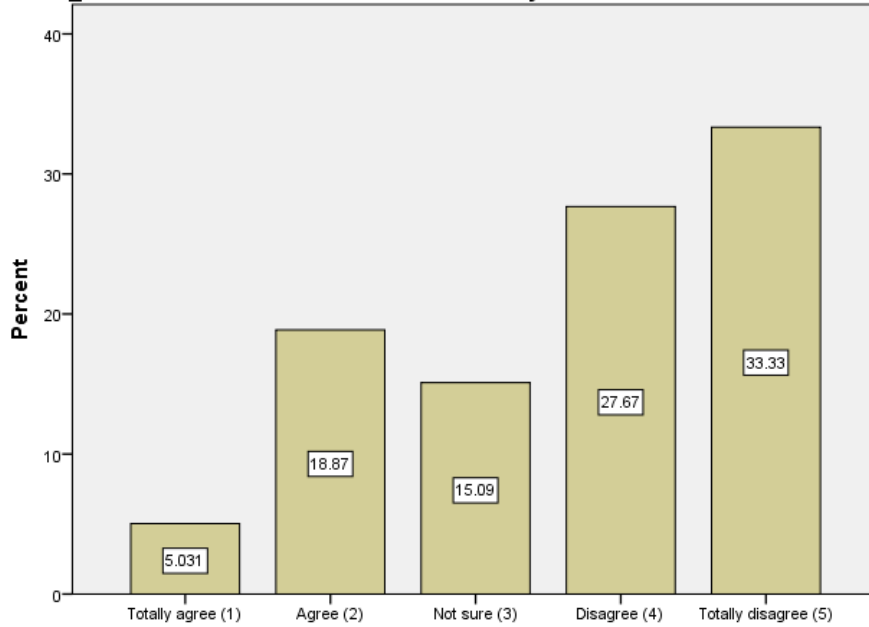
I don't use shredder for discarding the documents with sensitive information

Q2_24: I don't use shredder for discarding the documents with sensitive information



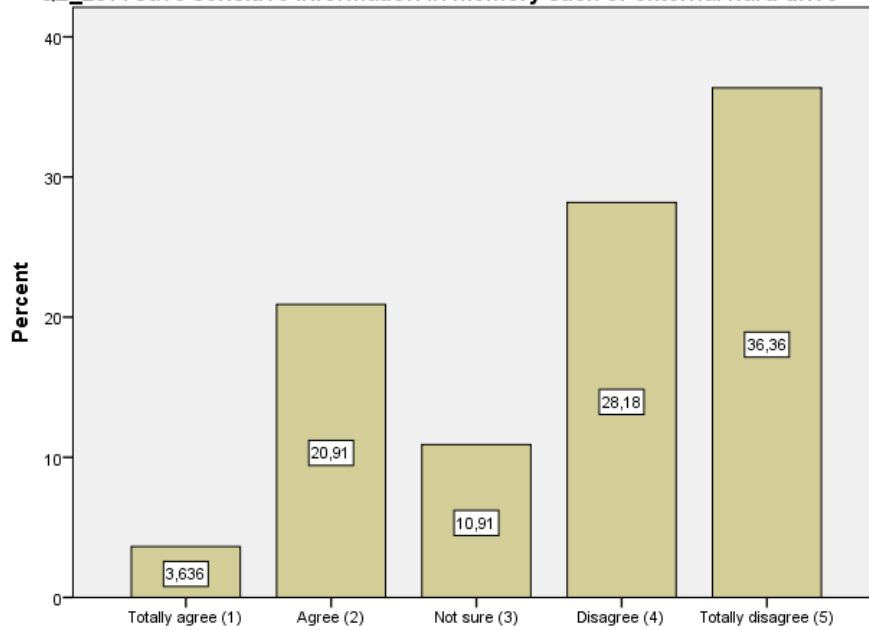
I don't use shredder for discarding the documents with sensitive information

Q1_22: I save sensitive information in memory stick or external hard drive

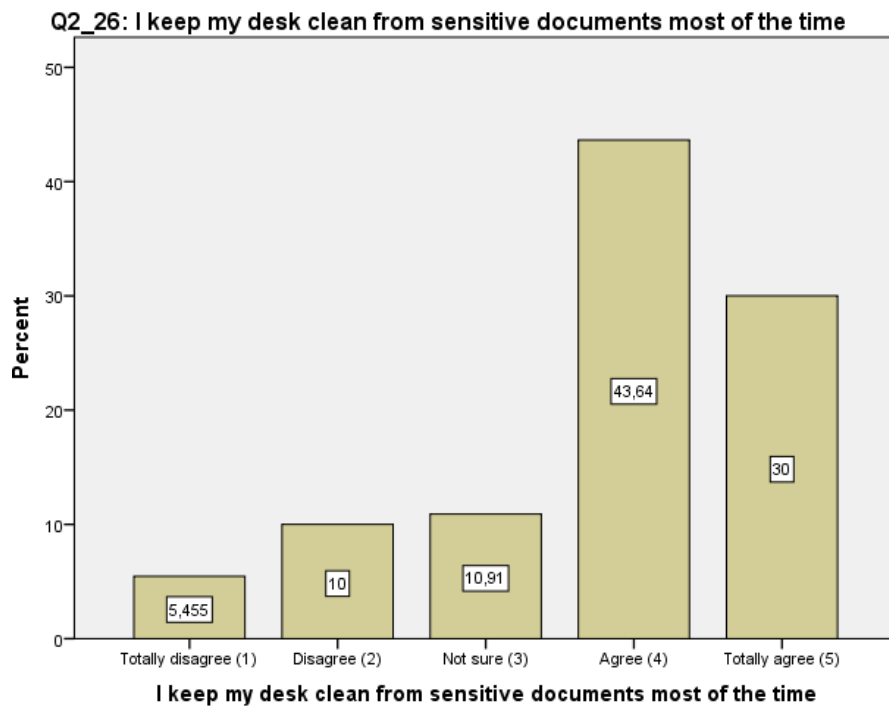
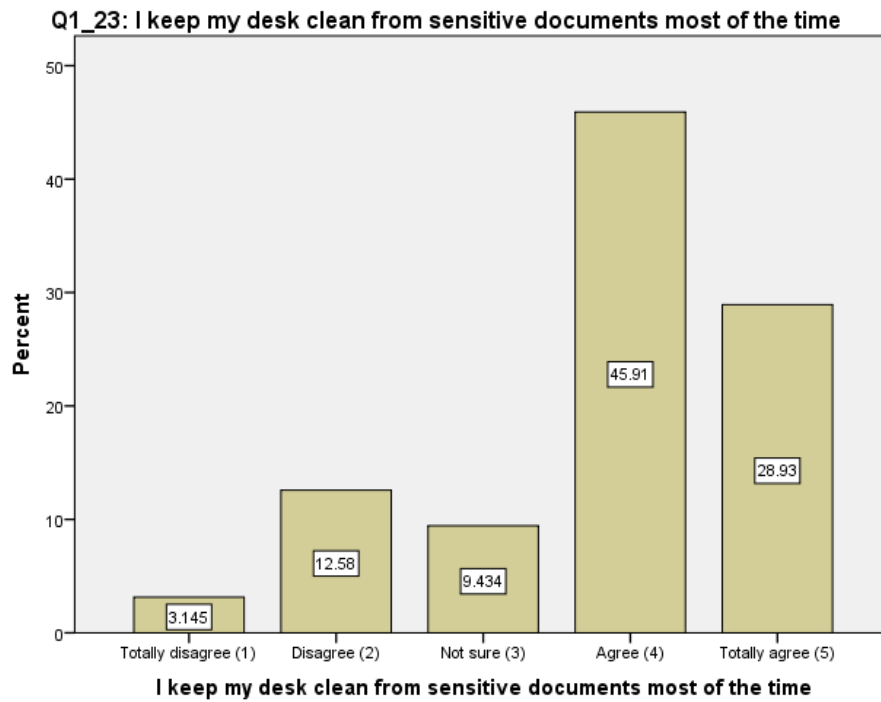


I save sensitive information in memory stick or external hard drive

Q2_25: I save sensitive information in memory stick or external hard drive



I save sensitive information in memory stick or external hard drive



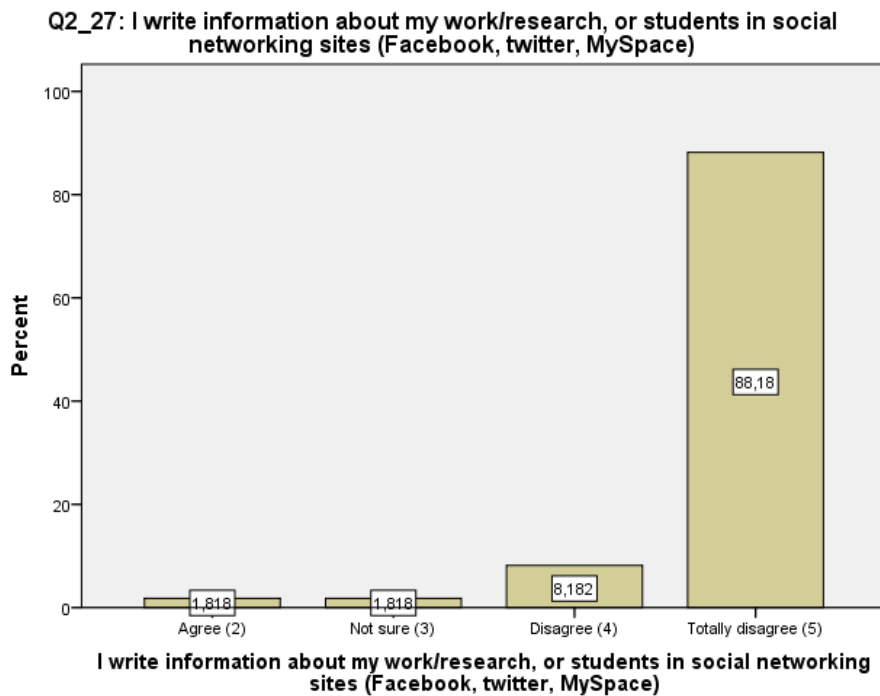
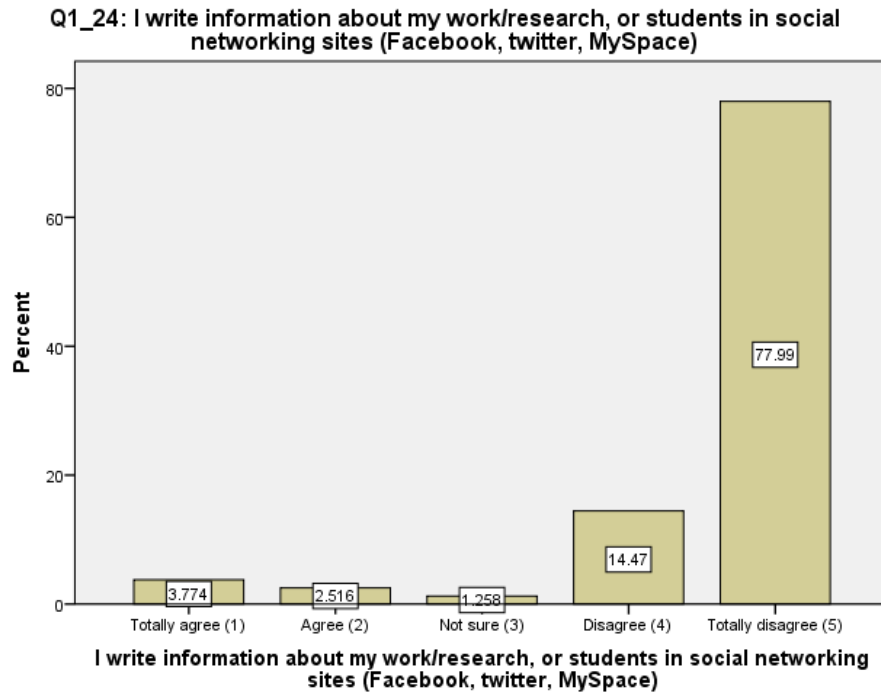
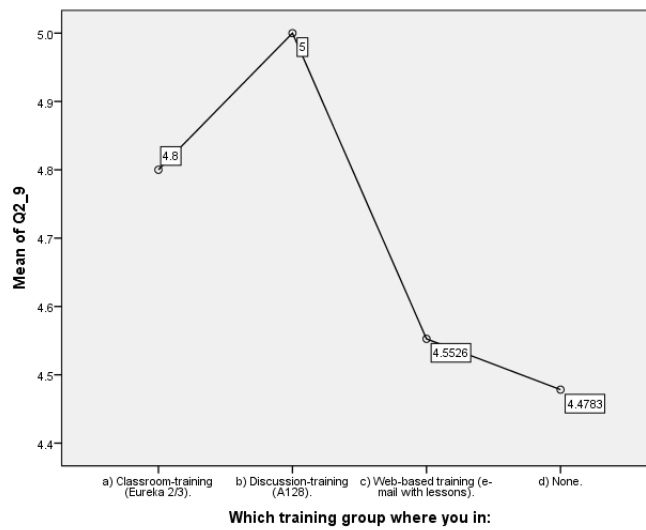
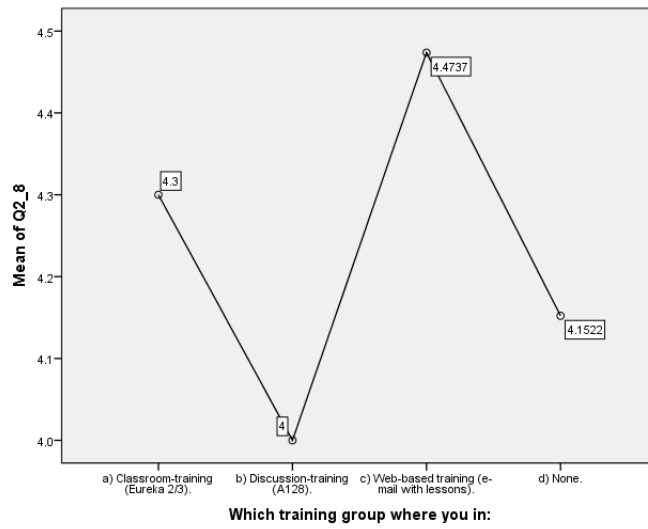
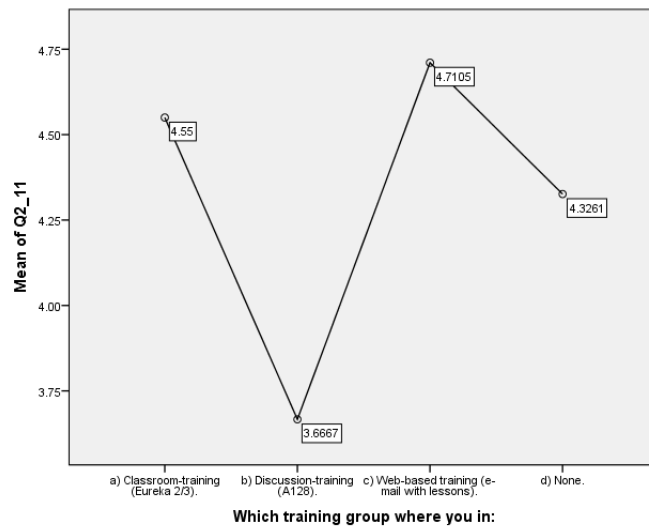
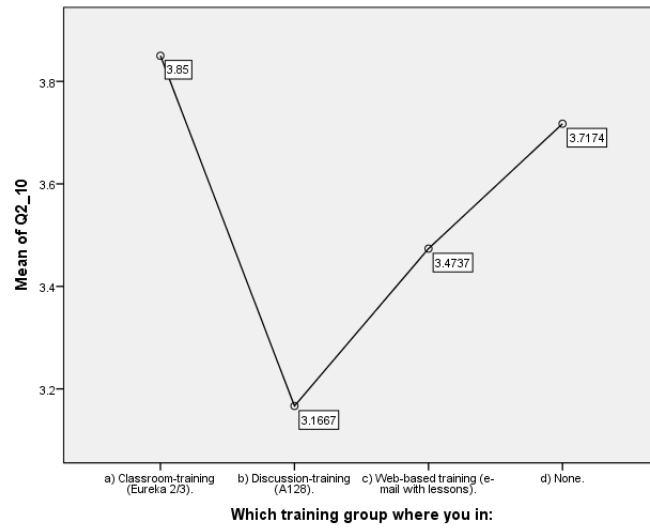


Figure 25: Second group - Policy Regulations Awareness.

C Mean Plots





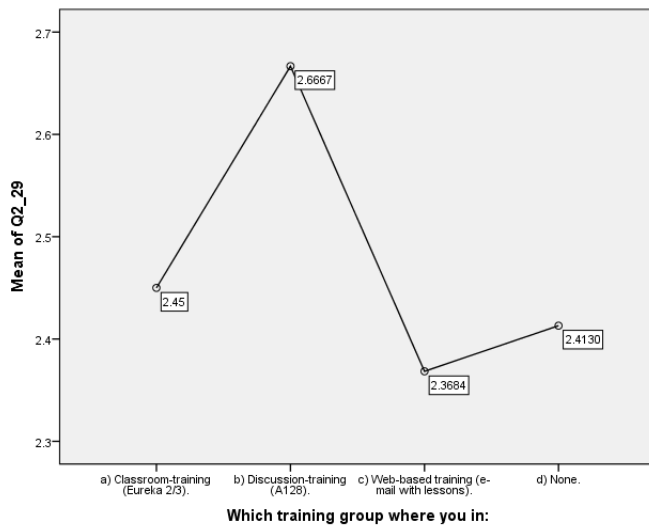
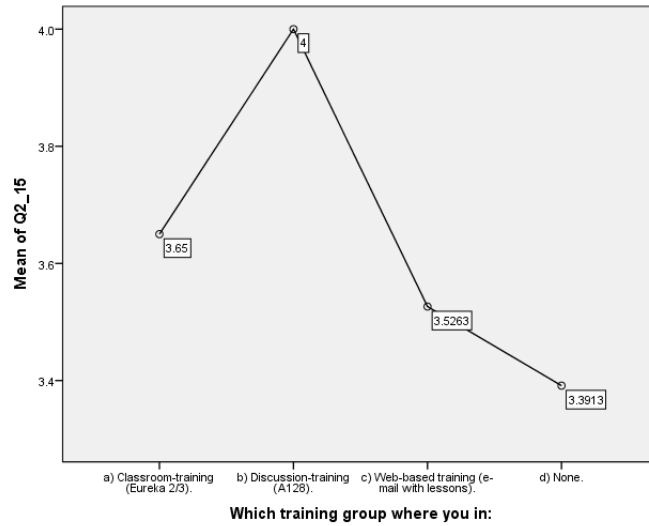
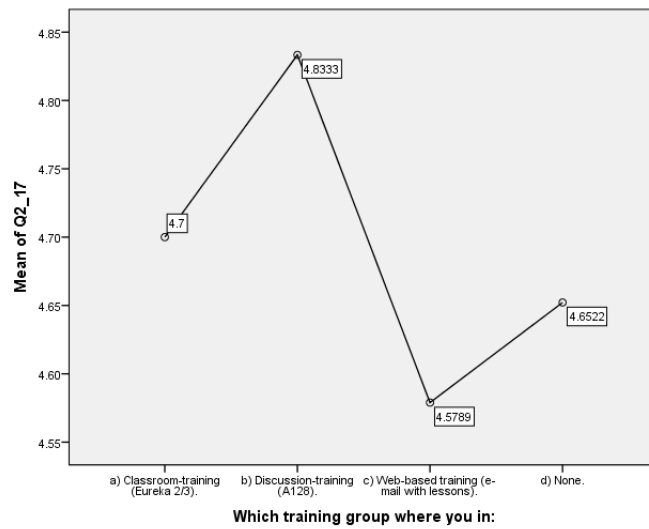
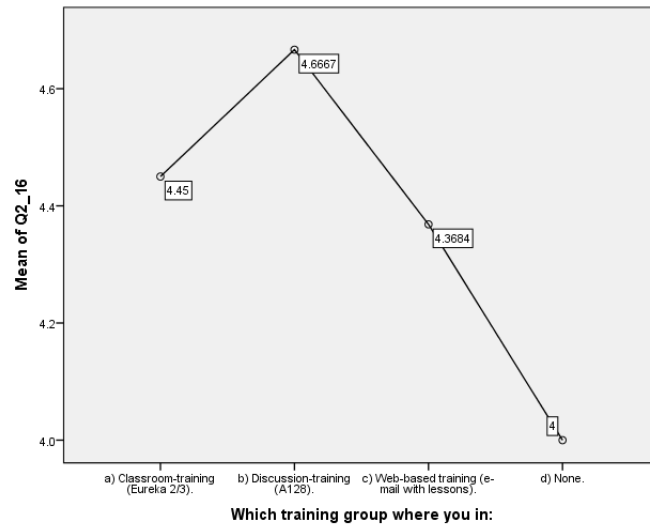
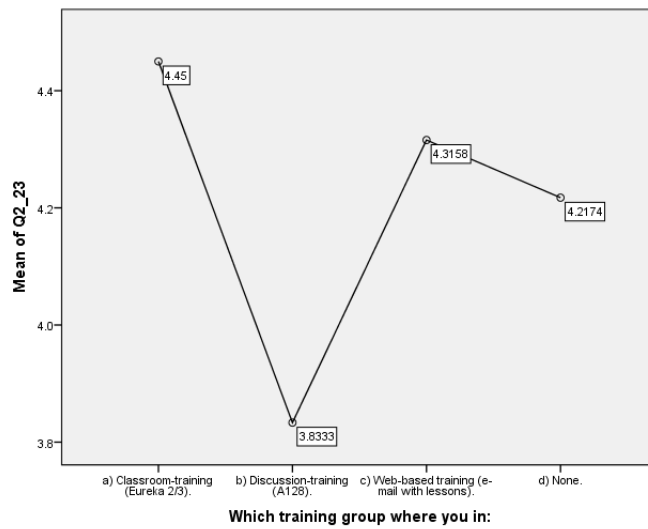
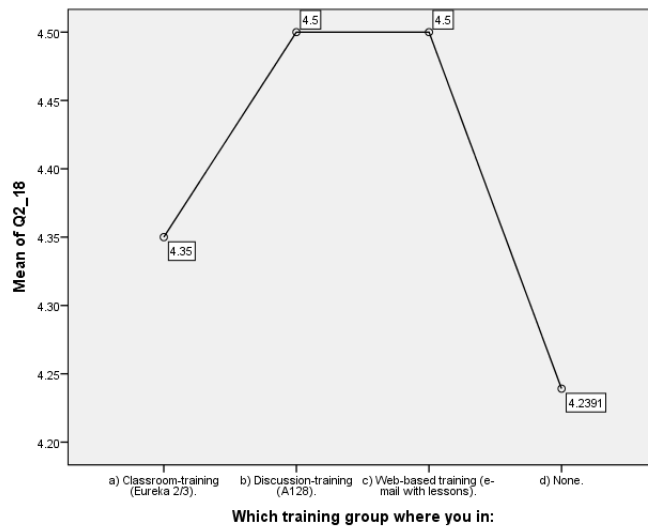
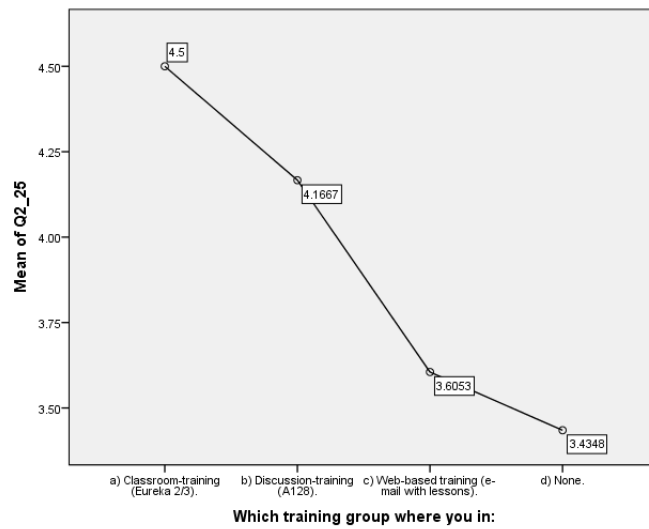
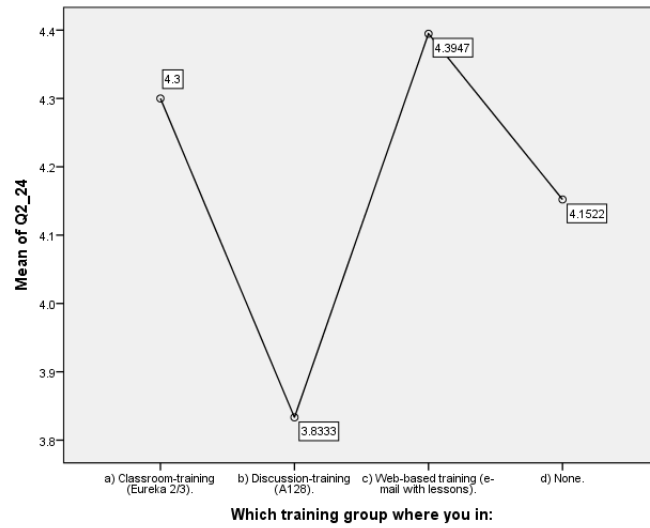


Figure 26: First group of questions (survey #2) - mean plots.







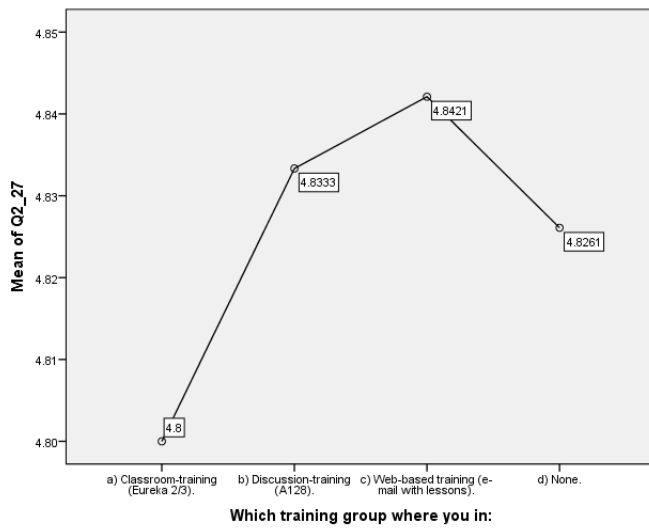
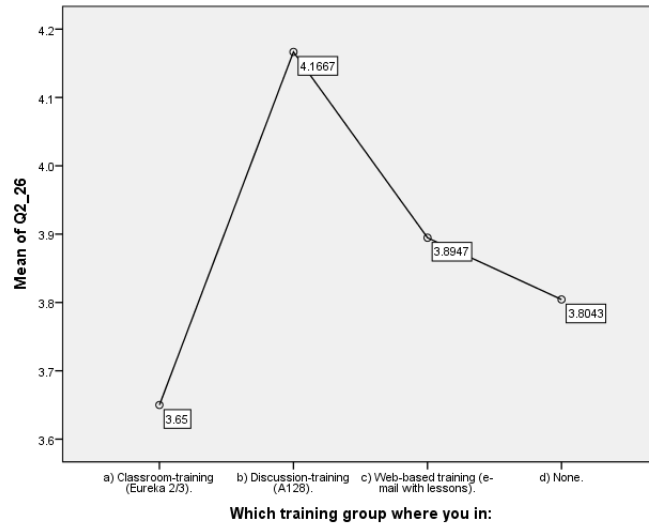


Figure 27: Second group of questions (survey #2) - mean plots.