UNIVERSITY OF TARTU

FACULTY OF MATHEMATICS AND COMPUTER SCIENCE

Institute of Software Engineering

Huishi Yin

Survey on Exploratory Software Testing

Master thesis (30 ECTS)

Supervisor: Dietmar Pfahl, University of Tartu

Author: .... "........" January 2014

Supervisor: .... "........" January 2014

Approved for defense

Professor: .... "........" January 2014

Tartu, 2014

Acknowledgments

Abstract

Exploratory Testing (ET) is a testing method that was created by Cem Kaner in 1983 [1]. The objective of this thesis is to investigate to what extent ET is currently applied in Estonian and Finnish software companies and what experience the companies have with using ET. Also, we want to know more about the relationship between role, working experience, location, company size, company history, type of test organization and the way how ET is used. In addition, we try to find out how testers think about ET, e.g., the advantages and disadvantages of ET, and how they plan to improve ET.

To achieve the objective of this thesis, we decided to use a survey to gather replies from testing relevant roles. After over 2 months, we got 61 replies. 50 are from Estonia and Finland. We summarize all information and answer three main research questions of this study, which are:

RQ 1: What characterizes those respondents that apply ET in Estonia and Finland and what experience do their companies have with using ET?
RQ 2: Which factors have an influence on using ET in Estonian and Finnish companies?
RQ 3: How do software engineers think about ET?

Respondents in Estonian and Finnish companies who apply ET are mostly testers (48%) and test managers (36%). Close to 50% of the respondents have more than five years of working experiences. More Estonian testers (50%) use ET than Finnish testers (38%). ET has been applied in small companies (less than 50 employees, 56%) relatively more often than in large companies (32%). Over 50% of the companies applying ET are older than 5 years. About 76% of the respondents who use ET are members of a separate testing organization within their company.

"Working experience" seems to be the main factor that influences whether a company applies ET in Estonia and Finland. Respondents who have much working experience use ET more than those who have less working experience. "Company size" is significantly different between those companies which apply ET in Estonia as compared to Finland. Respondents who are from smaller companies in Estonia prefer using ET, but in Finland, bigger companies' respondents use ET more than those who are from smaller companies.

Respondents consider that ET supports creativity, and ET are efficient and effective. Meanwhile some respondents think ET requires higher testing ability for testers, which means ET does not suit for everyone. Some respondents can use ET for so limit testing situations that make them considering that ET is inflexible. The main problem which ET brings is "hard to record". Most respondents are willing to change the recording strategy of ET.

Table of Contents

1. Introduction

1.1 Background

Testing is an important part of the software development life cycle. Testing helps assess correctness, completeness and quality of software in whole or in part. It is also an important sub-domain of SQA (software quality assurance). Testers try to use many different ways to find more defects. Exploratory Testing (ET) is a manual testing method that was first presented by Cem Kaner in 1983 [1]. He defines ET as "A style of software testing that emphasizes the personal freedom and responsibility of the individual tester to continually optimize the quality of his/her work by treating test-related learning, test design, test execution, and test result interpretation as mutually supportive activities that run in parallel throughout the project." [2]. Unfortunately, the definition of ET by Kaner is not very intelligible. In fact, after reading this long definition, testers might find the definition too abstract and maybe even confusing.

Along with the developing of testing technology, people have different descriptions and understandings of ET. Some may believe ET is 'ad hoc' testing while others may consider ET is a method of error guessing [3]. Some may even believe that ET means to do testing again and again [9]. The most simple and formal definition of ET was given by James Bach: "Exploratory testing is simultaneous learning, test design, and test execution. [8]"

1.2 About the Thesis

In this thesis, we don't attempt an in-depth research about defining ET or how to carry out specific ET progress. We are interested in finding out to what extent ET is currently applied in Estonian and Finnish software companies and what software engineers think about ET. Thus, the objective of this thesis is to investigate the characteristics of those software companies that apply ET in Estonia and Finland and what experience the companies have with using ET. Also, we want to know which factors favor using ET in a company, for example, the relationship between roles and using ET, working experiences and using ET, and so on. In addition, we try to find how software engineers think about ET, like advantage, disadvantage of ET and how to improve ET. Based on the research results, we want to see which understanding of ET is more predominant in practice.

1.3 Research Questions

Given the objectives of the thesis, three main research questions (RQ) should be answered. Each RQ has been extended into several specific questions.

RQ 1: What characterizes those respondents that apply ET in Estonia and Finland and what experience do their companies have with using ET?
    RQ 1.1 What is the distribution of respondents applying ET between Estonia and Finland?
    RQ 1.2 How big are the companies of respondents in Estonia and Finland who apply ET?

RQ 1.3 How old are the companies of respondents in Estonia and Finland who apply ET?

RQ 1.4 What kinds of software do software engineers test when using ET?

RQ 1.5 When do software engineers typically use ET?

RQ 1.6 Do software engineers use tools to support ET?

RQ 1.7 In what testing context do software engineers use ET?

RQ 2: Which factors have an influence on using ET in Estonian and Finnish companies?

RQ 2.1 Do testers use ET more frequently than other roles?

RQ 2.2 Do software engineers with more working experience use ET more frequently?

RQ 2.3 Does the location (Estonia or Finland) have an effect on applying ET?

RQ 2.4 Do software engineers in larger companies use ET more often than in smaller companies?

RQ 2.5 Is ET used more frequently in older companies?

RQ 2.6 Do software engineers in separate test organizations use ET more frequently?

RQ 2.7 What factors are different when software engineers apply ET between Estonia and Finland?

RQ 3: How do software engineers think about ET?

RQ 3.1 What elements consider software engineers as essential for defining ET?

RQ 3.2 What characteristics do software engineers think ET has?

RQ 3.3 What do software engineers think are advantages and disadvantages of ET?

RQ 3.4 Do software engineers want to improve ET, and how?

1.4 Research Method

To achieve the objective of this thesis, we designed a questionnaire which was made accessible online during a two-months period. The survey URL was promoted among the members of professional testers' associations in Estonia and Finland as well as a convenience sample of individuals in companies in both countries who promoted the survey within their companies. We also sent the link of our online survey system via mail lists, posted the link to forums, social networks and personal web blogs. The survey items can be found in **Appendix A**.

The reason why we used a survey for data collection is that we were interested in getting a broad overview. The results of the survey will be used as a reference point for follow-up research in the form of case studies. These case studies, however, are not in the scope of this thesis and will be conducted by others.

After the data was collected, we summarized and analyzed the data using descriptive and test statistics. For data collected in the form of open questions, we conducted a simple qualitative analysis. The related data and graphs can be found in **Appendix B.**

1.5 Thesis Outline

As well as presenting the main objective of this thesis, the first chapter lists the main research questions of the thesis and briefly summarizes the research method. Chapter 2, the progress of survey designing, improving and executing are described. In chapter 3, we describe how we summarize the data and the main findings of statistical analysis. Chapter 4 presents the compare of our study conclusions with other relevant research results and find out what we should do for further study. Chapter 5 shows our further plan for improving our study.

2. Survey

A survey is a "comprehensive system for collecting information to describe, compare or explain knowledge, attitudes and behavior" [6]. In the following we briefly describe the design of our survey and how the survey questions relate to the research questions listed in Section 1.

2.1 Survey Design

Before we design the survey, we should plan how to implement main research questions to specific questions, but first of all, we should set the object of the survey. Of cause, the three main research questions are the objectives of our survey. To answer those questions, we consider the survey should contain three parts:

Part 1: Questions on general background
In part 1, we think about questions helping us understand the answers about RQ 1 of this study. Questions about the general information which also the factors we are interested in can affect using ET should be asked. We should ask respondents their roles, working experiences, department location, company size, age of company, and software characteristics. We also want to ask them if there is a separate testing organization in their company, and whether they are one of the separate testing organizations.

Part 2: Questions on Exploratory Testing (ET)
In part 2, some aspects of RQ 1 and the RQs 2 and 3 should be answered. We want to know if respondents use ET, then "when, how, why" they use it, and how they understand ET. We also want to know what characteristics of ET make respondents like or dislike using ET, and how they want to improve ET. But if the respondents do not use ET, then we wonder which testing methods they use more frequently.

Part 3: Follow-up activities
In part 3, we explain the purpose of the survey, the use of the data, and the confidentiality of the respondents' information. We also ask respondents to leave their contact details voluntarily for sharing study results and further information collecting.

Based on the plan of designing the survey, we formulated 18 survey questions (SQ). Figure 2.1 shows the logic of the survey (part 1 and part 2). Table 2.1 shows how RQs (including sub-questions) relate to the 18 SQs.
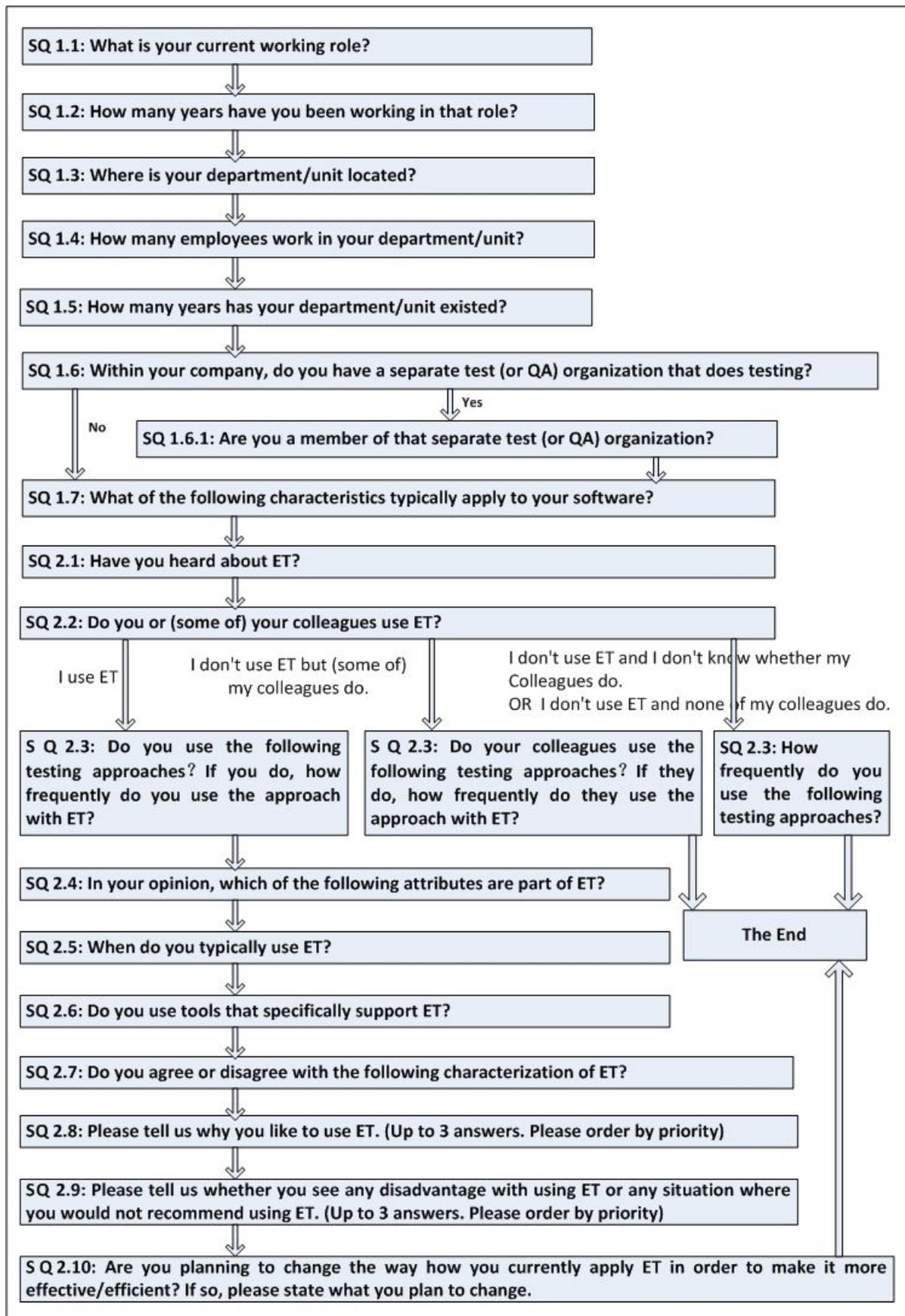
Figure 2.1 Logic of the survey

Research Question (RQ)	Survey Question (SQ)
RQ 1.1 What is the distribution of respondents applying ET between Estonia and Finland?	SQ 1.3: Where is your department/unit located? SQ 2.2: Do you or (some of) your colleagues use ET?
RQ 1.2 How big are the companies of respondents in Estonia and Finland who apply ET?	SQ 1.4: How many employees work in your department/unit? SQ 2.2: Do you or (some of) your colleagues use ET?
RQ 1.3 How old are the companies of respondents in Estonia and Finland who apply ET?	SQ 1.5: How many years has your department /unit existed? SQ 2.2: Do you or (some of) your colleagues use ET?
RQ 1.4 What kinds of software do software engineers test when using ET?	SQ 1.7: What of the following characteristics typically apply to your software?
RQ 1.5 When do software engineers typically use ET?	SQ 2.5: When do you typically use ET?
RQ 1.6 Do software engineers use tools to support ET?	SQ 2.6: Do you use tools that specifically support ET?
RQ 1.7 In what testing context do software engineers use ET?	SQ 2.3(a): Do you use the following testing approaches? If you do, how frequently do you use the approach with ET?
RQ 2.1 Do testers use ET more frequently than other roles?	SQ 1.1: What is your current working role? SQ 2.2: Do you or (some of) your colleagues use ET?
RQ 2.2 Do software engineers with more working experience use ET more frequently?	SQ 1.2: How many years have you been working in that role? SQ 2.2: Do you or (some of) your colleagues use ET?
RQ 2.3 Does the location (Estonia or Finland) have an effect on applying ET?	SQ 1.3: Where is your department/unit located? SQ 2.2: Do you or (some of) your colleagues use ET?
RQ 2.4 Do software engineers in larger companies use ET more often than in smaller companies?	SQ 1.4: How many employees work in your department/unit? SQ 2.2: Do you or (some of) your colleagues use ET?
RQ 2.5 Is ET used more frequently in older companies?	SQ 1.5: How many years has your department/unit existed? SQ 2.2: Do you or (some of) your colleagues use ET?
RQ 2.6 Do software engineers in separate test organizations use ET more frequently?	SQ 1.6.1: Are you a member of that separate test (or QA) organization? SQ 2.2: Do you or (some of) your colleagues use ET?
RQ 2.7 What factors are different when software	SQ 1.1, SQ 1.2, SQ 1.4, SQ 1.5, SQ 1.6.1

engineers apply ET between Estonia and Finland?	SQ 2.2: Do you or (some of) your colleagues use ET?
RQ 3.1 What elements consider software engineers as essential for defining ET?	SQ 2.4: In your opinion, which of the following attributes are part of ET?
RQ 3.2 What characteristics do software engineers think ET has?	SQ 2.7: Do you agree or disagree with the following characterization of ET?
RQ 3.3 What do software engineers think are advantages and disadvantages of ET?	SQ 2.8: Please tell us why you like to use ET. SQ 2.9: Please tell us whether you see any disadvantage with using ET or any situation where you would not recommend using ET.
RQ 3.4 Do software engineers want to improve ET, and how?	SQ 2.10: Are you planning to change the way how you currently apply ET in order to make it more effective/efficient? If so, please state what you plan to change.

Table 2.1 Relationship between RQ and SQ

2.2 Survey Implementation and Distribution Channels

From 10th of June to 31st of August, we were running the survey on the online survey system "Diaochapai" [4]. We posted the survey link to several testing relevant forums, social networks and personal web blogs. At the end of August, we closed the survey service. In total, we received 61 complete responses. The distribution of visit resources is shown in Figure 2.2 below. We can see from Figure 2.2 that mail lists ("Direct visit") were the best way to promote our survey (47.54%). Linkedin was also a good choice (22.95%).
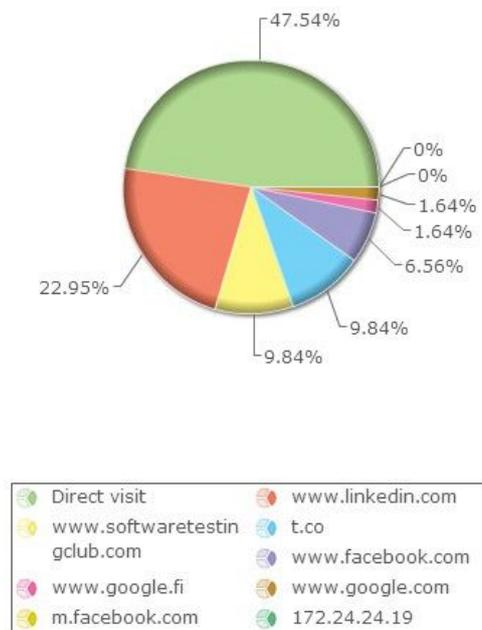
Figure 2.2 Visit resources

3. Data Analysis

3.1 Demographic Information

In the 61 received responses, we identified 11 responses that are neither from Finland nor Estonia. Thus, we decided to split the datasets into 2 groups. Response set 1, labeled "Esi & Fin" dataset contains 50 responses exclusively from Finland and Estonia. Response set 2, labeled "All" dataset, contains all 61 responses. Now we first give an overview of all data we have. The figure in the graph includes value and percentage. (e.g. 7, 11%, means 7 respondents choose this option and the percentage is 11%)
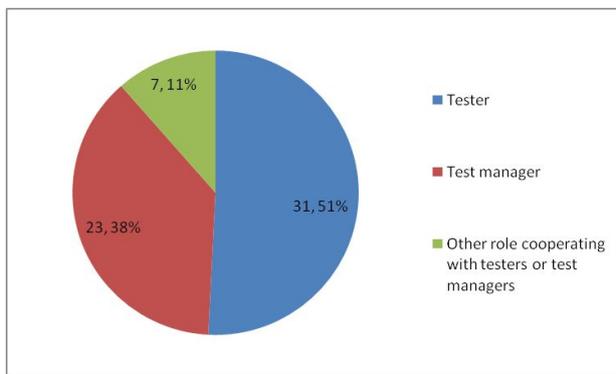
Figure 3.1 Respondents' role
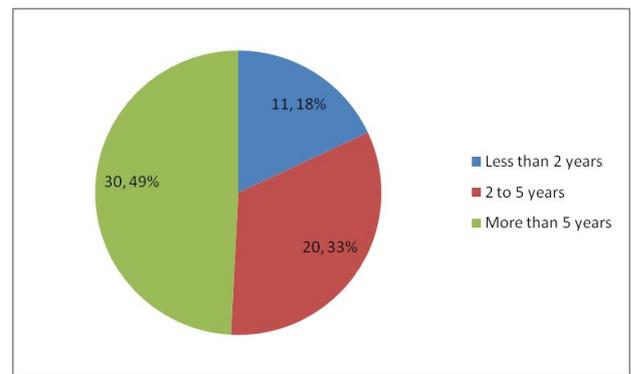
Figure 3.2 Respondents' working experiences
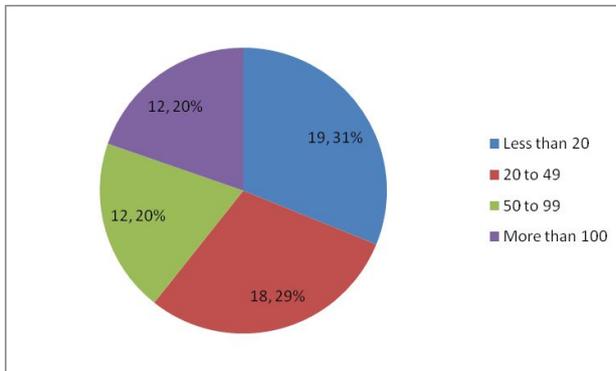
Figure 3.3 Respondents' company size
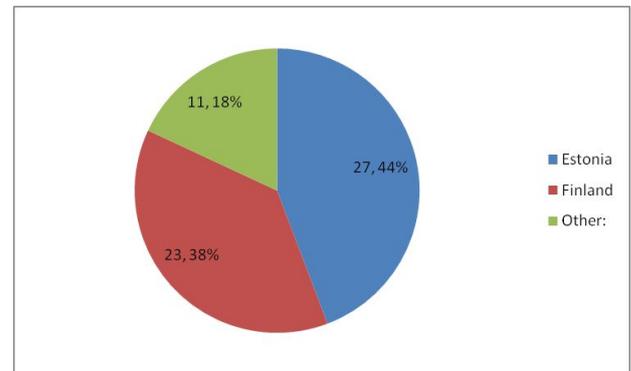
Figure 3.4 Respondents' company location
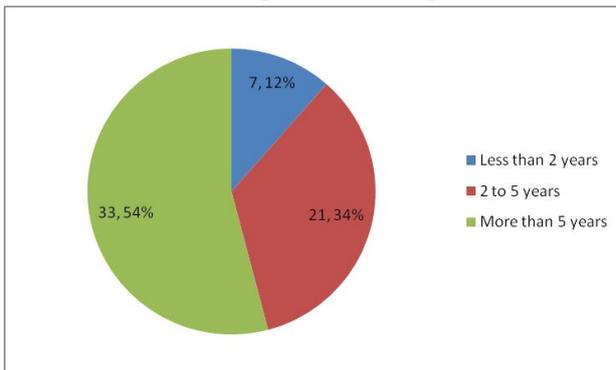
Figure 3.5 Age of respondents' company

Figure 3.1 shows the distribution of respondents' roles. 51% respondents are testers, 38% are tester managers, and other roles account for 11%. So we can see most respondents are testers or test managers (54 respondents, 89%).

Figure 3.2 shows 49% respondents have more than 5 years working experience in their role, and 22 respondents have 2 to 5 years working experience (33%). Less than one fifth of the respondents (18%) work less than 2 years in their role.

Figure 3.3 tells us, smaller (less than 50 employees) companies account for a higher proportion (less than 20 employees 31% and 20 to 49 employees 29%). Companies with over 50 employees account for 40% of all responses.

We can see form Figure 3.4 that most respondents are from Estonian software companies (27 respondents, 44%). Over one third (23 respondents, 38%) are from Finland and 11 respondents (18%) from other location which is unexpected.

In the last figure, Figure 3.5 shows over a half (33 respondents, 54%) of the companies have been created more than 5 years ago. 34% (21 respondents) of the respondents'companies have been created between 5 and 2 years ago. Only seven respondents (12%) are from a startup companies having less than 2 years company history.

3.2 Main Findings

This section shows the survey results relevant to each research question (RQ). The discussion will be put in next section.

3.2.1 Research Question 1

RQ 1: What characterizes those respondents that apply ET in Estonia and Finland and what experience do their companies have with using ET?

We received 61 complete responses, 27 are from Estonia, 23 are from Finland, and 11 are from other countries. Among the Estonian respondents, 25 use ET and two say they do not use ET (and they do not know whether their colleagues in their company use ET). Among the Finnish respondents 19 say they use ET and one sais they know someone in their company who uses ET, three say they do not use ET and they do not know whether their colleagues in their company use ET.

To answer RQ1.1, RQ1.2, RQ1.3, we first analyze the responses to SQ1.3, SQ 1.4, and SQ1.5, respectively. Since we are only interested in the data from Estonia and Finland, we remove the 11 responses from other locations. Then we used the tool SPSS to analyze the data showing the relations between two SQs' result. For instance, if we want to know how many respondents from Estonia are using ET, we should know two groups of data (how many respondents are from Estonia and how many of those are using ET). We need to know the responses from SQ1.3 and SQ2.2 and find out the mixed part of those results. To make

the result clearer, we merge the four options of SQ2.2 into two groups, "Yes" and "No". We consider those respondents who did not take "Yes" as their response, no matter their "colleagues" use ET or not, the response of this question is "No".

RQ 1.1 *What is the distribution of respondents applying ET between Estonia and Finland?*

	Use ET	Do not use ET	Sum
Estonia	25 (92.59%)	2 (7.41%)	27 (54%)
Finland	19 (82.61%)	4 (17.39%)	23 (46%)
Sum	44 (88%)	6 (12%)	50 (100%)

Cross-table 3.1 Location & using ET

In order to answer RQ1.1, we applied the cross table analysis using SPSS. For example, cross-table 3.1 combines responses to SQ1.3 and SQ2.2 to find out how many respondents use ET in Estonia and how many respondents use ET in Finland. Cross-table 3.1 shows that 50 respondents are from Estonia or Finland. 54% (27 respondents) are from "Estonia" and 46% (23 respondents) are from "Finland". After analyzing with SQ2.2, 92.59% of 27 respondents are from Estonia and also use ET, The corresponding percentage for respondents from Finland is 82.61% of 23 respondents. The distribution of respondents applying ET is 25 respondents (92.59% * 54% = 50% of all respondents) use ET in Estonia and 19 respondents (82.61% * 82.61% = 38% of all respondents) use ET in Finland.

RQ 1.2 *How big are the companies of respondents in Estonia and Finland who apply ET?*

	Use ET	Do not use ET	Sum
Less than 20	14 (93.33%)	1 (6.67%)	15 (30%)
20 to 49	14 (87.5%)	2 (12.5%)	16 (32%)
50 to 99	8 (80%)	2 (20%)	10 (20%)
More than 100	8 (88.89%)	1 (11.11%)	9 (18%)
Sum	44 (88%)	6 (12%)	50 (100%)

Cross-table 3.2 Company size & using ET

Cross-table 3.2 describes 15 (30%) respondents are from software companies which have less than 20 employees and 14 respondents from these companies say they use ET. 16 (32%) are from those which have more than 20 but less than 49 employees and14 respondents (87.5%) from these companies say they use ET. 19 (38%) are from companies which have more than 50 employees and 16 respondents claim that they use ET. We can see from cross-table 3.2 that ET has been applied in smaller size companies (less than 50 employees,

28 respondents) more often than in bigger companies (more than 50 employees, 16 respondents).

RQ 1.3 How old are the companies of respondents in Estonia and Finland who apply ET?

	Use ET	Do not use ET	Sum
Less than 2 years	3 (75%)	1 (25%)	4 (8%)
2 to 5 years	16 (84.21%)	3 (15.79%)	19 (38%)
More than 5 years	25 (92.59%)	2 (7.41%)	27 (54%)
Sum	44 (88%)	6 (12%)	50 (100%)

Cross-table 3.3 Company age & using ET

Cross-table 3.3 shows that 25 respondents (50%) are from software companies which apply ET and have been created more than 5 years ago. 16 (34%) respondents are from those companies which have been created less than 5 but more than 2 years ago. Only three (6%) respondents are from startup companies having less than 2 years company history. We can see that more respondents who use ET are working for older software companies.

To answer RQ1.4 to RQ1.7, we analyze the responses to SQ1.7, SQ2.5 and SQ 2.6, respectively. When answering RQs 1.4 to 1.7, we use all 51 datasets of respondents saying they are using ET, i.e., including responses from all locations (also from outside Estonia and Finland).

RQ 1.4 What kinds of software do software engineers test when using ET?
To answer RQ 1.4, we analyze the responses to SQ1.7. SQ 1.7 offers four main characteristics that respondents applying ET can check to characterize the software to which they apply ET. Table 3.1 summarizes the frequencies of characterization patterns chosen by the respondents. 14 respondents said their software has all four characteristics, and four of the respondents chose the option "none of above". Two respondents chose three characteristics excluding "It is usability-critical". Three respondents only excluded option "It has high security demand". Two respondents only excluded option "It is performance-critical". Eight respondents only excluded option "It is safety-critical". Four respondents chose "It is usability-critical" and "It has high security demand". Six respondents say their software is "usability-critical" and "performance-critical". Only one respondent selected options "It has high security demand" and "It is safety-critical". Seven respondents selected only "It is usability-critical". One respondent selected only "It has high security demand". Two respondents selected only "It is performance-critical". Most software tested with ET has all characteristics. Only four respondents use ET to test software without any of these characteristics. 44 respondents use ET when their software is "usability-critical". 35 respondents use ET when their software is "performance-critical". 32 respondents use ET when their software has "high security demand". 22 respondents use ET when their software

is "safety-critical". We cannot say these four characteristics are the factors that necessarily should be tested by ET, but we can state that respondents use ET to test software having these characteristics.

Characteristic													Sum
It is usability-critical	√		√	√	√	√	√		√				44 (32.1%)
It has high security demand	√	√		√	√	√		√		√			32 (23.4%)
It is performance-critical	√	√	√		√		√				√		35 (25.5%)
It is safety-critical	√	√	√	√			√						22 (16.1%)
None of above												√	4 (2.9%)
Number of selected	14	2	3	2	8	4	6	1	7	1	2	4	137

Table 3.1 characteristics of respondents' software

RQ 1.5 *When do software engineers typically use ET?*

Our hypothesis was twofold: If software engineers use ET as one of the first test activities, it may be because software engineers want to know their testing object better; If software engineers use ET as one of the last test activities, it may be because software engineers consider ET can find defects which cannot be found by other test methods. To answer RQ1.5, we analyze the responses to *SQ2.5: When do you typically use ET?* Ten respondents (19%) use ET as one of the first test activities. Only five respondents (9%) use ET as one of the last test activities. 39 respondents (72%) use ET at any time during testing. We can see from the result, most respondents use ET at any time during testing, It may be because respondents think ET is a flexible testing approach which can be used at any time.

RQ 1.6 *Do software engineers use tools to support ET?*

To answer RQ 1.6, we analyze responses to *SQ 2.6.* 15 respondents say they use tools to support ET and they list what tools they use. 39 respondents (72%) said they do not use tools. Table 3.2 shows the frequency of tools mentioned by respondents when applying ET. We can see from the result that most respondents say they use software to support ET. In this case, "Mind Maps" and "Rapid Reporter" are the most used tools. Some software engineers prefer non-software tools, like literature and check list etc. Three respondents use special tools (Custom made tool) to support ET. The fact that most respondents do not use tools to support ET shows tools are not necessary needed for supporting ET. Even though when respondents use tools to support ET, normal tools are used more frequently. Special tools for ET are not a popular choice.

Type of tools	Tools	Frequency
Software	Mind Maps (e.g. Xmind)	7
	Rapid Reporter	5
	Custom made tool	3
	Evernote	2

	Excel	1
	qTrace	1
	Microsoft Test Manager	1
	Vim-Editor	1
	Jira Test Sessions	1
	OneNote	1
	Perlclip	1
	IETester	1
	BB Flashback	1
	Screen Recorder	1
Non-software	Literature	2
	PostIts	1
	Check lists	1
	Paper & pen	1
	Total	32

Table 3.2 Special tools with frequency of usage

***RQ 1.7** In what testing context do software engineers use ET?*
SQ2.2 and SQ2.3(a) are designed for getting information about the testing context in which software engineers like to use ET. We offered ten common testing approaches as options in SQ2.3(a). Figure 3.6 shows the summed-up responses of SQ2.3(a). Since we offered ten different test approaches, we got 540 responses, i.e., ten from each of the 54 respondents who said they are using ET. Each response states whether an approach is being used by the respondent, and if so, whether it is used with ET always, often, sometimes or never.
Among the 540 responses, 166 (30.8%) responses are "use approach with ET always". System Testing is the most frequently used approach with ET (28 respondents said they always do system testing with ET). 25 respondents use Smoke Testing always with ET.

56 (10.3%) responses are "never use ET" with any approach we offered. Automated Testing is the approach most frequently never used with ET, 21 respondents say they never use it with ET. 12 respondents never use Performance Testing with ET.

81 (15%) responses are "use approach with ET sometimes" and 112 (20.7%) responses are "often use approach with ET". As a result, we got 359 (66.5%) responses stating that the testing approach we offered is used with ET always, often or sometimes. The result indicates

that ET can be used in many testing contexts, even in combination with automated testing and performance testing.

Since most respondents using ET don't do unit testing (40 out of 54), this is the approach that is also the least used with ET.
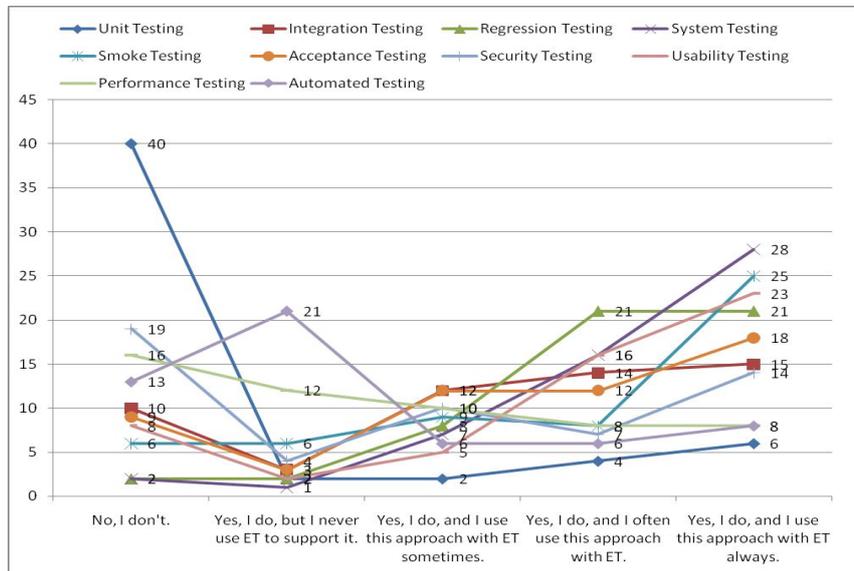
Figure 3.6 Testing approaches and frequency of usage with ET

3.2.2 Research Question 2

RQ 2 Which factors have an influence on using ET in Estonian and Finnish companies?
To answer RQ2.1 to RQ2.6, we use both groups of datasets, i.e., "Esi & Fin" dataset (results can be found in this section) and "All" dataset (compared with the "Esi & Fin" dataset, results can be found in Appendix C). We conduct statistical analysis. To answer RQ2.1 to RQ2.7, we first set α=0.05. Because the dataset is qualitative data, we use the Chi-square test [7]. When the dataset is over 40 but there are more than 25% of cells that have expected number less than 5, then we should use the Exact Chi-square testing method to do the test [8] and compare the "Fisher's Exact Test" value (P value) with α. We use statistic software SPSS to do analysis steps as below:
1) Set hypothesis H0 and H1.
2) Do cross-table analysis by using question results.
3) Choose accurate "Chi-square testing" method to do the test.
4) Judge whether H0 or H1 should be accepted based on the test results.

RQ 2.1 Do testers use ET more frequently than other roles?
Our hypothesis was there are no significant differences in the frequency of using ET for different roles.
1) Set hypothesis H0 and H1.
    H0: There are no differences in the frequency of using ET for different roles.
    H1: There are differences in the frequency of using ET for different roles.

2) Do cross-table analysis by using question results of SQ1.1 and SQ2.2.
3) Choose "Exact Chi-square testing" method to do the test.
4) Compare "Fisher's Exact Test" value (P value) with α. Judge whether H0 or H1 should be accepted based on the test results.

Role * Reply Crosstabulation

Count

		Reply		Total
		Yes	No	
Role	Tester	24	2	26
	Tester Manager	18	2	20
	Other roles	2	2	4
Total		44	6	50

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	6.002[a]	2	.050	.086		
Likelihood Ratio	4.042	2	.133	.128		
Fisher's Exact Test	4.610			.103		
Linear-by-Linear Association	3.183[b]	1	.074	.093	.078	.058
N of Valid Cases	50					

a. 4 cells (66.7%) have expected count less than 5. The minimum expected count is .48.
b. The standardized statistic is 1.784.

Table 3.3 Role & Reply Cross-tabulation

From Table 3.3, we can see P=0.103> 0.05, so accept H$_0$, which means there are no significant differences between roles in using ET. After double-checking with "All" dataset group (see Appendix C.1 RQ2.1), we get the same result.

RQ 2.2 Do software engineers with more testing experience use ET more frequently?
Our hypothesis was that there is no significant difference in using ET for different levels of working experience.
1) Set hypothesis H0 and H1.
    H0: There is no correlation in using ET for different levels of working experience.
    H1: There is a correlation in using ET for different levels of working experience.
2) Do cross-table analysis by using question results of SQ1.2 and SQ2.2.
3) Choose "Exact Chi-square testing" method to do the test.
4) Compare "Fisher's Exact Test" value (P value) with α. Judge whether H0 or H1 should be accepted based on the test results.

From Table 3.4, we can see, P=0.016< 0.05, so reject H0, accept H1, which means there is a significant correlation in using ET for different levels of working experience. The higher the working experience the more is ET used. After double-checking with "All" dataset group (see Appendix C.1 RQ2.2), we got a different result, i.e., H0 is accepted. But the P value from "All" dataset is 0.063, which is quite close to 0.05.

Year * Reply Crosstabulation

Count

		Reply		Total
		Yes	No	
Year	< 2	7	3	10
	2-5	14	3	17
	> 5	23	0	23
Total		44	6	50

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	6.718[a]	2	.035	.036		
Likelihood Ratio	8.631	2	.013	.022		
Fisher's Exact Test	7.004			.016		
Linear-by-Linear Association	6.515[b]	1	.011	.013	.013	.010
N of Valid Cases	50					

a. 3 cells (50.0%) have expected count less than 5. The minimum expected count is 1.20.
b. The standardized statistic is -2.552.

Table 3.4 Working experience (years) & Reply Cross-tabulation

***RQ 2.3** Does the location (Estonia or Finland) have an effect on applying ET?*
Our hypothesis was that there is no significant difference in using ET for different geographical locations (i.e., Estonia versus Finland).
1) Set hypothesis H0 and H1.
    H0: There is no difference in using ET for different geographical locations.
    H1: There is a difference in using ET for different geographical locations.
2) Do cross-table analysis by using question results of SQ1.3 and SQ2.2.
3) Choose "Exact Chi-square testing" method to do the test.
4) Compare "Fisher's Exact Test" value (P value) with α. Judge whether H0 or H1 should be accepted based on the test results.

Location * Reply Crosstabulation

Count

		Reply		Total
		Yes	No	
Location	Estonia	25	2	27
	Finland	19	4	23
Total		44	6	50

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	1.172[a]	1	.279	.395	.259	
Continuity Correction[b]	.418	1	.518			
Likelihood Ratio	1.180	1	.277	.395	.259	
Fisher's Exact Test				.395	.259	
Linear-by-Linear Association	1.149[c]	1	.284	.395	.259	.196
N of Valid Cases	50					

a. 2 cells (50.0%) have expected count less than 5. The minimum expected count is 2.76.
b. Computed only for a 2x2 table
c. The standardized statistic is 1.072.

Table 3.5 Company location & Reply Cross-tabulation

From Table 3.5, we can see, P=0.395 > 0.05, accept H0. There is no significant difference in in using ET for different geographical locations. After double-checking with "All" dataset group (see Appendix C.1 RQ2.3), we get the same result.

***RQ 2.4** Do software engineers in larger companies use ET more often than in smaller companies?*

Our hypothesis was there is no significant difference in using ET for different sizes of the respondent's company.

1) Set hypothesis H0 and H1.

H0: There is no difference in using ET for different company sizes.

H1: There is a difference in using ET for different company sizes.

2) Do cross-table analysis by using question results of SQ1.4 and SQ2.2.

3) Choose "Exact Chi-square testing" method to do the test.

4) Compare "Fisher's Exact Test" value (P value) with α. Judge whether H0 or H1 should be accepted based on the test results.

From Table 3.6, we can see, P=0.881 > 0.05, accept H0. There is no significant difference in using ET for different sizes of the respondent's company. After double-checking with "All" dataset group (see Appendix C.1 RQ2.4), we get the same result.

Company size * Reply Crosstabulation

Count

		Reply		Total
		Yes	No	
Number of employees	≤ 20	14	1	15
	20-49	14	2	16
	50-99	8	2	10
	> 100	8	1	9
Total		44	6	50

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	1.021[a]	3	.796	.929		
Likelihood Ratio	1.001	3	.801	.929		
Fisher's Exact Test	1.326			.881		
Linear-by-Linear Association	.334[b]	1	.563	.694	.350	.130
N of Valid Cases	50					

a. 4 cells (50.0%) have expected count less than 5. The minimum expected count is 1.08.
b. The standardized statistic is .578.

Table 3.6 Company size & Reply Cross-tabulation

RQ 2.5 *Is ET used more frequently in older companies?*

Our hypothesis was there is no significant difference in using ET for different ages of the respondent's company.

1) Set hypothesis H0 and H1.

H0: There is no difference in using ET for different company ages.

H1: There is a difference in using ET for different company ages.

2) Do cross-table analysis by using question results of SQ1.5 and SQ2.2.

3) Choose "Exact Chi-square testing" method to do the test.

4) Compare "Fisher's Exact Test" value (P value) with α. Judge whether H0 or H1 should be accepted based on the test results.

From Table 3.7, we can see, P=0.410 > 0.05, accept H0. There is no significant difference in using ET for different company ages. After double-checking with "All" dataset group (see Appendix C.1 RQ2.5), we get the same result.

Company history * Reply Crosstabulation

Count

		Reply		
		Yes	No	Total
Company history (years)	< 2	3	1	4
	2-5	16	3	19
	> 5	25	2	27
Total		44	6	50

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	1.438[a]	2	.487	.507		
Likelihood Ratio	1.361	2	.506	.507		
Fisher's Exact Test	2.022			.410		
Linear-by-Linear Association	1.408[b]	1	.235	.310	.195	.126
N of Valid Cases	50					

a. 4 cells (66.7%) have expected count less than 5. The minimum expected count is .48.
b. The standardized statistic is -1.186.

Table 3.7 Company age (years) & Reply Cross-tabulation

RQ 2.6 *Do software engineers in separate test organizations use ET more frequently?*

Our hypothesis was there is no significant difference in using ET no matter whether there is a separate test organization or not in the respondent's company.

1) Set hypothesis H0 and H1.

H0: There is no difference in using ET no matter whether there is a separate test organization or not.

H1: There is a difference in using depending on whether there is a separate test organization or not.

2) Do cross-table analysis by using question results of SQ1.6.1 and SQ2.2.

3) Choose "Exact Chi-square testing" method to do the test.

4) Compare "Fisher's Exact Test" value (P value) with α. Judge whether H0 or H1 should be accepted based on the test results. After double-checking with "All" dataset group (see Appendix C.1 RQ2.6), we get the same result.

Separate organization * Reply Crosstabulation

Count

		Reply		
		Yes	No	Total
Separate organization	Have	22	2	24
	Don't have	3	2	5
Total		25	4	29

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	3.490[a]	1	.062	.127	.127	
Continuity Correction[b]	1.335	1	.248			
Likelihood Ratio	2.771	1	.096	.127	.127	
Fisher's Exact Test				.127	.127	
Linear-by-Linear Association	3.369[c]	1	.066	.127	.127	.116
N of Valid Cases	29					

a. 3 cells (75.0%) have expected count less than 5. The minimum expected count is .69.
b. Computed only for a 2x2 table
c. The standardized statistic is 1.836.

Table 3.8 Separate organization & Reply Cross-tabulation

From Table 3.8, we can see, P=0.127 > 0.05, accept H0. There is no significant difference in using ET no matter whether there is a separate test organization or not in the respondent's company.

After double-checking with "All" dataset group (see Appendix C.1 RQ2.6), we get the same result.

RQ 2.7 What factors are different when software engineers apply ET between Estonia and Finland?

To answer this question, we did the same analyses as described above for RQ2.1 to RQ2.6, but using only a sub-set of the "Esi & Fin" dataset. Because we only want to know what factors are different between Estonia and Finland when using ET, we removed those datasets where respondents said they did not use ET. Then we classified the data by location. After analyzing with SPSS, we got the following tables. Only Table3.9 shows P value equals 0.008, less than 0.05, which means that respondents who said they used ET work in companies of different size when comparing between Estonia and Finland. (Find other analysis results in Appendix C.2.)

Company size * Reply Crosstabulation

Count

		Reply		Total
		Estonia	Finland	
Size	< 20	12	2	14
	20-49	9	5	14
	50-99	2	6	8
	> 100	2	6	8
Total		25	19	44

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	11.685[a]	3	.009	.006		
Likelihood Ratio	12.449	3	.006	.011		
Fisher's Exact Test	11.442			.008		
Linear-by-Linear Association	10.508[b]	1	.001	.001	.001	.001
N of Valid Cases	44					

a. 4 cells (50.0%) have expected count less than 5. The minimum expected count is 3.45.
b. The standardized statistic is 3.242.

Table 3.9 Company size & reply cross-tabulation

3.2.3 Research Question 3

RQ 3: How do software engineers think about ET?

Considering our limited amount of data, we think that using all 61 datasets is a better way to address the research question. Of cause, there are 54 datasets of those using ET. To answer RQ3.1 to RQ3.4, we analyze SQ2.4, SQ2.7, SQ2.8, SQ2.9 and SQ2.10. We use "All" dataset as analysis resource.

RQ 3.1 What elements consider software engineers as essential for defining ET?

In this question, we designed a question based on the concept of Session-Based Test Management (SBTM), which James Bach described as "A method for measuring and managing exploratory testing" [10]. In SQ2.4, we give nine options related to the SBTM concept. The option "ET has a defect log" had the most advocates (77.78%) and option "ET has a test log" stands the second place (over 70%). Our hypothesis was that software engineers would not be able to distinguish between ET and SBTM. However, only two respondents chose option "None of above" (3.7%). In the "Other" option, three respondents

wrote "do not mix ET with SBTM". Given the data, we thinks that most respondents mix SBTM with ET. This may have two reasons. First, respondents do not have a clear definition of ET. Second, respondents may consider SBTM is a necessary part of ET. Especially "defect log" and "test log" was relatively often checked which shows that respondents think logging is very important for ET.
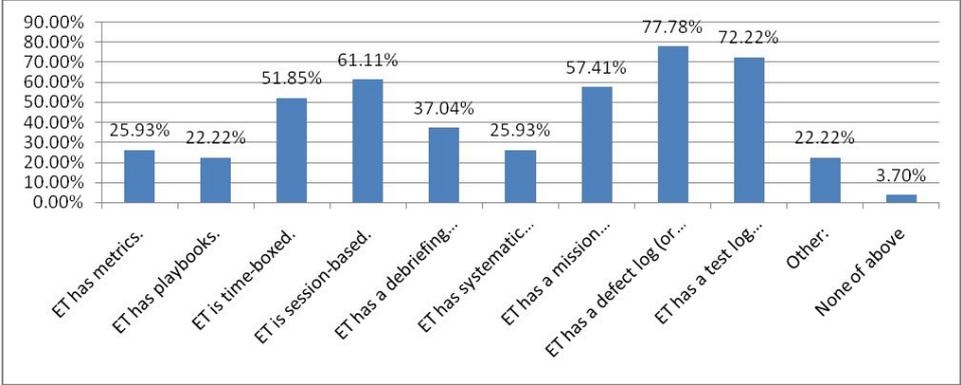
Figure 3.7 Result of SQ2.4

RQ 3.2 *What characteristics do software engineers think ET has?*

We offered seven option to characterize ET. For each option, the median is 27 (54 times 50%). Thus, we checked whether the option of each item in cumulative frequency is equal to or above 27 [5]. From Figure 3.8 we can see that respondents chose "Agree" or "Strongly agree" for each option we gave more frequently than 27 times (vertical line in Figure 3.8). Especially, respondents strongly agree that "ET supports creativity", "ET makes testing interesting and engaging" and "ET is flexible".
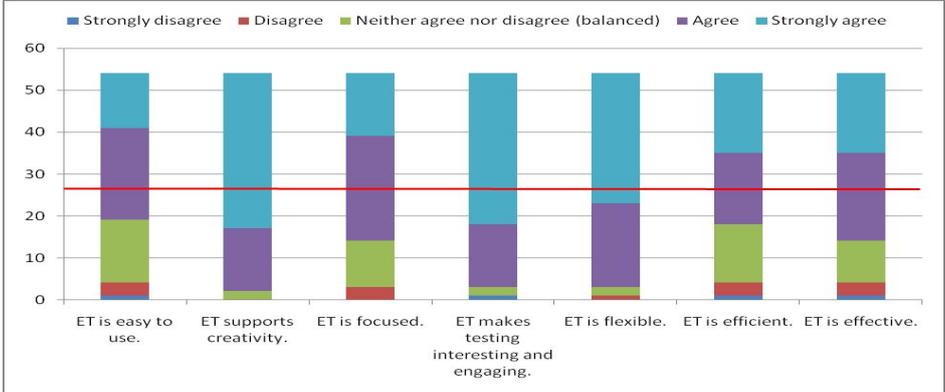
Figure 3.8 Cumulative frequencies of ET characteristics

RQ 3.3 *What do software engineers think are advantages and disadvantages of ET?*

To answer RQ3.3, we asked software engineers to list 3 reasons why they like and dislike (SQ2.9 and SQ2.9) to use ET (Advantages and Disadvantages). We assume that answers were ranked, i.e., the first mentioned reason has a higher weight than later ones. We first draw the key words from respondents' answer. Then set weight 3, 2, 1 to reason No.1 to No.3 and then times frequency (Frequency 1 means the frequency of Advantage No.1; Score 1 equals Frequency 1 times weight), and sum up all the scores. The highest score ranks the top. We can see from Table 3.10 and Table 3.11, the top 3 advantages are: Supports

creativity, Efficient, Effective. The top 3 disadvantages are: High requirement for tester, Inflexible, Hard to record.

Advantage	Frequency 1	Score 1	Frequency 2	Score 2	Frequency 3	Score 3	Total
Supports creativity	6	18	12	24	5	5	47
Efficient	10	30	6	12	1	1	43
Effective	7	21	3	6	5	5	32
Flexible	3	9	3	6	2	2	17
Supports learning	3	9	3	6	2	2	17
Time saving	4	12	1	2	1	1	15
Interesting	2	6	3	6	3	3	15
Easy	3	9	2	4	0	0	13
Emphasizes tester	2	6	1	2	1	1	9
Focused	1	3	0	0	1	1	4
Essential	1	3	0	0	0	0	3
Independent	0	0	1	2	0	0	2
Clear data	0	0	0	0	1	1	1
Create logs	0	0	0	0	1	1	1
Total	42	126	35	70	23	23	

Table 3.10 Rank of Advantages

Disadvantage	Frequency 1	Score 1	Frequency 2	Score 2	Frequency 3	Score 3	Total
high requirement for tester	9	27	4	8	1	1	36
inflexible	5	15	0	0	1	1	16
hard to record	2	6	4	8	0	0	14
not good for complicated project	4	12	0	0	0	0	12
not all-inclusive	2	6	0	0	0	0	6
time consuming	2	6	0	0	0	0	6
no focus	1	3	1	2			5

confusing	1	3	0	0	0	0	3
hard to compare results	1	3	0	0	0	0	3
inefficient	1	3	0	0	0	0	3
stakeholders don't appreciate	1	3	0	0	0	0	3
time limit	1	3	0	0	1	1	4
too popular	1	3	1	2	0	0	5
unnecessary	1	3	0	0	0	0	3
unrepeatable	1	3	0	0	0	0	3
hard to report	0	0	4	8	0	0	8
ineffective	0	0	2	4	0	0	4
uncontrollable	0	0	2	4	0	0	4
energy consuming	0	0	1	2	0	0	2
inaccurate	0	0	1	2	0	0	2
stakeholders don't appreciate	0	0	1	2	0	0	2
Total	33	99	21	42	3	3	

Table 3.11 Rank of Disadvantages

RQ 3.4 Do software engineers want to improve ET, and how?
Table 3.12 shows the plan of changing ET offered by software engineers. There are six responses relevant to record and report of ET. Respondents think that "record" and "report" are at the top of the list of what should be improved of ET, which indicates that recording and reporting are very important for applying ET. ET should not be a free testing approach without any record or report.

Your plan for changing	Frequency
Create a record for ET	3
Improve report	2
Find a better reporting system for ET	1
More risk-based testing.	1
Study more and have more experience	1
Use a good tool	1

Use ET more often	1
Use SBTM and TBTM together	1
Choose the way of doing ET according to the project	1
Improve testing all the time	1
Do ET in the morning	1
Total	15

Table 3.12 Software engineers' plan for changing ET

3.3 Discussion

3.3.1 Applying ET in Estonia and Finland

Among the 61 responses, 50 respondents are from Estonia and Finland. Figure 3.9 shows the distribution of applying ET between Estonia and Finland (from the result of RQ1.1). There are 23 respondents (46%) from Finland and 27 respondents (54%) are from Estonia. 93% of respondents use ET in Finland and 83% respondents use ET in Estonia.
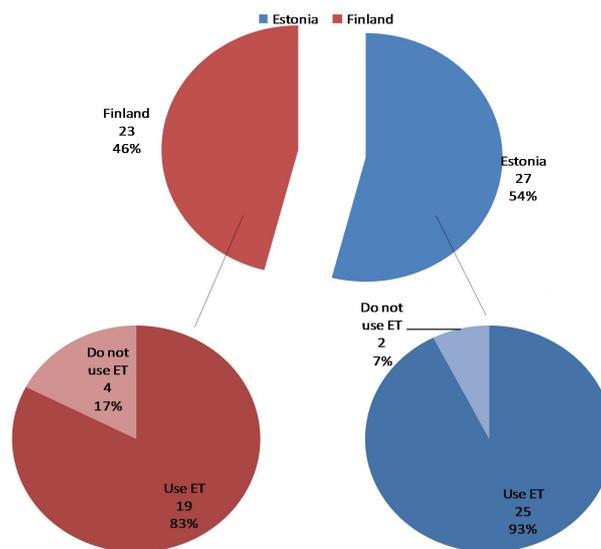
Figure 3.9 Distribution of applying ET between Estonia and Finland

Figure 3.10 and 3.11 show the distribution of applying ET in Finland and Estonia based on 4 different factors (from result of RQ2.7). We can see, in Finland and Estonia, the distribution of Role and Age of company are quite same. But respondents who have more than 5 years' working experiences in Finland have a higher percentage than in Estonia. Respondents who are from companies having more than 50 employees in Finland are more frequent than the number in Estonia. Instead, respondents from Estonia are more from companies having less than 50 employees.
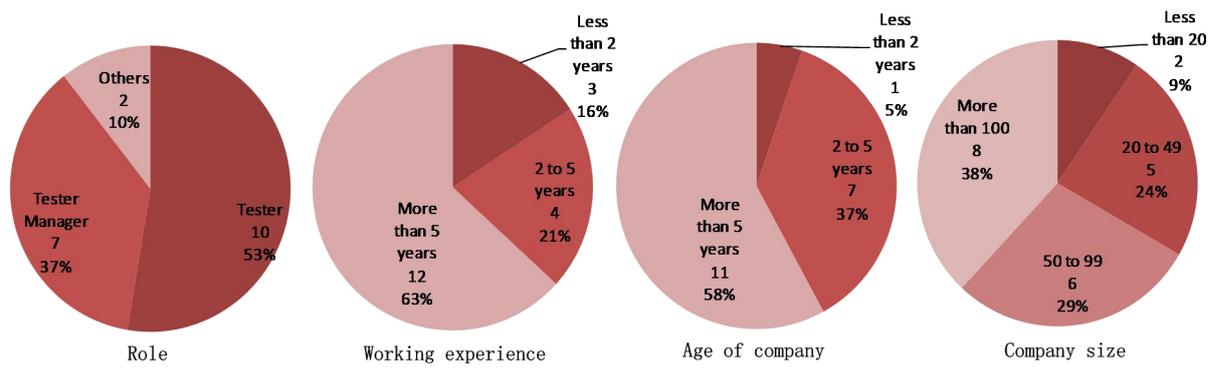
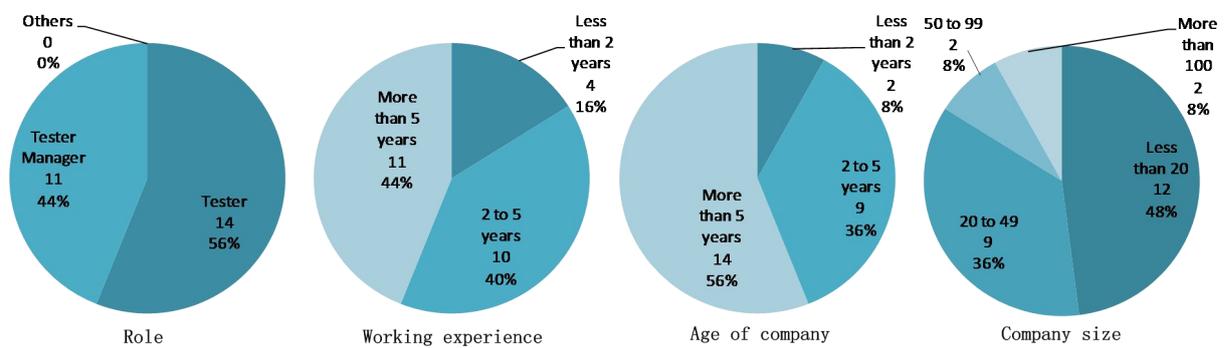Figure 3.10 Distribution of applying ET in Finland

Figure 3.11 Distribution of applying ET in Estonia

3.3.2 Experience of Applying ET

The result for RQ1.4 shows most software tested by ET has at least one of the characteristics we offered. Those characteristics we offered are very common characteristics of software, so we can say that most software could be tested using ET. The result for RQ1.5 shows most respondents use ET at any time during testing. The result for RQ1.6 shows very few respondents use special tools for testing ET, while most respondents say they do not use tools for testing ET, only a few respondents use normal tools to test ET, like Excel, Text, and so on. The result regarding to RQ1.7 is that many kinds of testing approaches are used with ET frequently. The results for RQ1.5 and RQ1.6 indicate that using ET is not restricted by the testing time and availability of tools. All findings from RQ1.4, RQ1.5, RQ1.6 and RQ1.7 verify our hypothesis that ET is a flexible approach. Can be used to test most software at any time during testing and special tools are not necessary for supporting ET, meanwhile ET is easy to implement with other testing approaches.

3.3.3 Factors Affecting the Usage of ET

From the results for RQ2.1 to RQ2.7, we know only "working experiences" relates to using ET in Estonia and Finland. Those respondents who have more working experience prefer using ET (from result of RQ2.2). Companies applying ET in Estonia and Finland have different size. In Estonia, respondents who are from smaller companies having less than 20 employees prefer ET, but in Finland, respondents who are from bigger companies having more than 20 employees like using ET more than smaller one.

3.3.4 Understanding of ET

The results for RQ3.1 show a very interesting mixed understanding of ET and SBTM. A large proportion of respondents consider the elements of SBTM are part of ET, which indicates most respondents do not understand the definition of ET and SBTM. But it also indicates SBTM is important for applying ET. We can see from the results for RQ3.2 that the evaluation of ET is very positive. Most respondents strongly agree that ET supports creativity and makes testing interesting and engaging and ET is effective and flexible. They also agree that ET is focused and ET is efficient and easy to use. It was partly same results shown in RQ3.3, which the top 3 advantages are: Supports creativity, Efficient, Effective. But there are 3 top disadvantages proposed as: High requirement for tester, Inflexible, Hard to record. Now we get interesting results. All characteristics affirmed in RQ3.2 are also mentioned in RQ3.3, not only in the advantage list, but also in the disadvantage list. For example, "inflexible", "inefficient", "no focus" and "ineffective" are mentioned in disadvantage list. Especially, inflexible is one of top three disadvantages. It shows these four characteristics are controversial, even though most respondents consider these characteristics are advantages of ET.

The results for RQ3.4 show "create a record" is the thing that respondents want to do for improving ET. It also can be seen in RQ3.2 that "hard to record" is the third disadvantage. RQ3.1shows most respondents think ET should have "defect or test log". In our study, one of conclusions would be no record of process and result is the most disadvantage of ET.

3.3.5 Threats to Validity

Each study has their shortages, no matter how we try to avoid it. Now we discuss the threats to validity of our study.

Selection bias

First of all, our survey has very strong pertinence. The title of our survey, "Survey on Exploratory Software Testing", shows our object is about ET. This may cause losing potential respondents at very beginning. For example, when a tester who does not know ET very well sees a survey about ET, he or she might skip the survey immediately. This may be also the reason why we got 100% positive responses (respondents heard about ET before) from SQ2.1. It's hard to prove that all software engineers heard about ET, even though the result is 100% positive. Secondly, although we declared the target respondents are those who work in Estonia and Finnish companies, we still got 11 out of 61 respondents from neither Estonia nor Finland. Meanwhile we would not get more respondents from Estonia and Finland because of limited diffusion of our survey.

External validity

The low number of responses would be anther weakness of our study, which is a big threat of external validity. There must be thousands of software engineers in Estonia and Finland and we received only 50 responses. Also, we do not know exactly how many of all the software

engineers we reached and what was the response rate. This makes our results hard to generalize to the population of software engineers located in Estonia and Finland.

Internal validity

The threats of internal validity are mainly from the design of the survey questions. The quality of survey depends on if the SQ formulates RQ accurately and the answer options cover all possibility. In other words, whether respondents can understand the SQs and answer the SQs easily. To reduce the threats of internal validity, we review and update all SQs for 5 times. In the sixth round, we sent our SQs to three experts in ET field. They gave lots of advices, including whether the SQs are readable and understandable. We improved our survey based on their advices to raise the internal validity of survey.

Statistical conclusion validity

Statistical conclusion validity and sample size, sample coverage has a strong relationship, the greater the number of samples, the more extensive coverage, and the more effective statistical conclusion. The biggest threat of statistical conclusion validity in our study is the sample size is not very high. In order to reduce the threat, we would like to work in the future, to collect more samples.

4. Related Work

Research about applying ET in Estonia and Finland is difficult to find, and actually there does not seem to exist any comparable surveys on the practice of ET in either Finland or Estonia. While I could not find any related paper reporting surveys similar to those reported in this thesis, I found out some discussion on the advantages and disadvantages of the ET. We concluded the advantages of ET are: "efficient", "efficiency", "easy to use with other testing approaches", and disadvantage is "no record". These results partly resemple those reported in [11]. We propose the disadvantage "high requirement for testers" and "inflexible" as new points which have not been mentioned in the literature on ET as of today.

5. Further Plan

In this study, we found some interesting answers on the practice of ET in Finland and Estonia. For instance, in the responses from SQ2.3(a) eight respondents say they use "Automated Testing" with ET. This is a quite interesting idea for us. Another interesting finding is that the results for SQ2.8 and SQ2.9 show some characteristics are proposed both in advantage and disadvantage list. As a result of the existence of these interesting answers, we would like to interview these respondents and try to find out why they answered like that. The follow-up interview questions (IQ) would be:

IQ1: Why don't you use ET? Do you think ET will bring problems to you and your company?

IQ2: Why you use Automated Testing with ET? And how do you implement?

IQ3: Why do you think ET is inflexible/ inefficient/ no focus/ ineffective? Do you have such bad experiences with ET?

We also notice that 61 responses is not a very good statistic amount. Implementing the survey again and trying to get more responses are in our further plan. We would like to have an agreement with some companies located in Estonia and Finland and make sure their employees answer our survey, so that we can get a higher response rate and a more representative set of responses.

6. Resümee

Ülevaade uurivast tarkvara testimisest

Magistritöö (30 ECTS)
Huishi Yin

Uuriv testimine on testimismeetod, millele pani aluse Cem Kaner 1983. aastal. Käesoleva magistritöö eesmärk on analüüsida, millises ulatuses kasutatakse uurivat testimist Eesti ja Soome tarkvaraettevõtetes ning milline on nende ettevõtete kogemus uuriva testimise osas. Samuti on eesmärgiks otsida seoseid rollide, töökogemuse, asukoha, ettevõtte suuruse, ettevõtte ajaloo, testimise korralduse ja uuriva testimise kasutamise viisi vahel. Lisaks pakub huvi, mida testijad arvavad uuriva testimise meetodist, millised on uuriva testimise eelised ja puudused ning kuidas plaanitakse uurivat testimist täiustada.

Eesmärgi saavutamiseks viidi läbi küsitlus isikute hulgas, kes on tööalaselt seotud testimisega. Kahe kuu jooksul saatsid vastused 61 küsitletut, kellest 50 on pärit Eestist ja Soomest. Töös tehakse kokkuvõte saadud vastustest ning vastatakse kolmele peamisele uurimisküsimusele, milleks on:
1) Mis iseloomustab küsitlusele vastanuid, kes kohaldavad uurivat testimist Eestis ja Soomes ja milline kogemus on nende ettevõtetel uuriva testimise kasutamisega?
2) Millised faktorid mõjutavad uuriva testimise kasutamist Eesti ja Soome ettevõtetes?
3) Mida arvavad tarkvaratehnikud uurivast testimisest?

Küsitlusele vastanute hulgas on kõige rohkem testijaid (48%) ja juhtivaid testijaid (36%). Ligi 50% vastanutest on rohkem kui 5 aastat töökogemust. Eesti testijatest kasutab uurivat testimist suurem osa (50%) kui Soome testijatest (38%). Uurivat testimist kohaldatakse väikestes ettevõtetes (vähem kui 50 töötajat, 56%) rohkem kui suurtes (32%). Üle 50% ettevõtetest, kus uurivat testimist kohaldatakse, on vanemad kui 5 aastat. Ligikaudu 76% vastanutest, kes kasutavad uurivat testimist, kuuluvad oma ettevõttes eraldiseisvasse testimisorganisatsiooni.

Töökogemus tundub olevat Eestis ja Soomes peamine faktor, mis mõjutab, kas uurivat testimist kasutatakse või mitte. Vastanud, kellel on pikaajaline töökogemus, kasutavad uurivat testimist rohkem kui need, kellel on väiksem töökogemus. Ettevõtte suurus on märgatavalt erinev Eesti ja Soome uurivat testimist kohaldavate ettevõtete vahel. Eestis eelistavad kasutada uurivat testimist vastanud väikestest ettevõtetest, kuid Soomes kasutavad uurivat testimist rohkem suurtest ettevõtetest pärit vastajad.

Vastajad toovad välja, et uuriv testimine soosib loovust ning on efektiivne. Samas arvavad mõned vastanud, et uuriv testimine nõuab testijalt suuremat testimisalast võimekust, mis tähendab, et uuriv testimine ei sobi igaühele. Mõned vastanud saavad kasutada uurivat testimist vaid vähestes olukordades, mistõttu leiavad nad, et uuriv testimine ei ole paindlik. Peamine probleem uuriva testimisega on, et seda on raske salvestada. Enamik vastanutest on valmis muutma uuriva testimise salvestamise strateegiat.

7. References

[1] C. Kaner, J. Falk, and H. Q. Nguyen, *Testing Computer Software (Second Edition)*, Van Nostrand Reinhold, New York, 1993, p 6, 7-11.

[2] C. Kaner, *A Tutorial in Exploratory Testing*, 2008, p 36. *http://www.kaner.com/pdfs/QAIExploring.pdf*[Online], Seen: 2013-10-05

[3] J. Itkonen, M. V. Mantyla, C. Lassenius, *Defect Detection Efficiency: Test Case Based vs. Exploratory Testing*, First International Symposium on Empirical Software Engineering and Measurement, 2007, P 2.

[4] "Diaochapai" online survey system. *http://www.diaochapai.com/intro*[Online]

[5] J. L. Wang, X. J, Liang, 定性数据分析 *(Qualitative Data Analysis)*, East China Normal University, Shanghai, 2005, P 3-4.

[6] S. L. Pfleeger, B. A. Kitchenham, *Principles of Survey Research Part 1 Turning Lemons into Lemonade*, Software Engineering Notes vol 26 no 6, 2001, p 16.

[7] Z. L. Deng, J. H. Sun, 定性数据分析方法在问卷调查中的简单应用 *(Qualitative data analysis methods in a simple application of the questionnaire)*, Shanghai Statistics, 2002, P 33

[8] J. Bach, Exploratory Testing, in The Testing Software engineer, Second ed., E. van Veenendaal Ed. Den Bosch: UTN Publishers, 2004.

[9] billliu, 探索性测试揭秘 *(The exploratory testing revelation)*, *http://blogs.msdn.com/b/billliu/archive/2012/05/26/10310656.aspx*[Oneline], Seen: 2013-12-11

[10] J. Bach, *Session-Based Test Management*, Software Testing and Quality Engineering magazine, 2000

[11] S. M. A. Shah, C. Gencel, U. S. Alvi, K. Petersen, *Towards a hybrid testing process unifying exploratory testing and scripted testing*, Journal Of Software: Evolution And Process, 2013,P 8.

8. Appendices

A. Survey Items

Category 1: Questions on general background

1.1: What is your current working role? * (Required, Multiple choice)

Answer categories:
- Tester
- Test manager
- Other role cooperating with testers or test managers

1.2: How many years have you been working in that role? * (Required, Multiple choice)

Answer categories:
- Less than 2 years
- 2 to 5 years
- More than 5 years

1.3: Where is your department/unit located? * (Required, Multiple choice)

Answer categories:
- Estonia
- Finland
- Other (open answer):

1.4: How many employees work in your department/unit? * (Required, Multiple choice)

Answer categories:
- Less than 20
- 20 to 49
- 50 to 99
- More than 100

1.5: How many years has your department/unit existed? * (Required, Multiple choice)

Answer categories:
- Less than 2 years
- 2 to 5 years
- More than 5 years

1.6: Within your company, do you have a separate test (or QA) organization that does testing? * (Required, Multiple choice)

Answer categories:
- Yes   (relevant to SQ1.6.1)
- No

1.6.1: Are you a member of that separate test (or QA) organization? * (Required, Multiple choice)

Answer categories:
- Yes
- No

1.7: What of the following characteristics typically apply to your software? * (Required, Check boxes)

Answer categories:
- It is usability-critical (e.g., it has a complex GUI which is important for the end user)
- It has high security demand.
- It is performance-critical.
- It is safety-critical.
- None of above.

Category 2: Questions on Exploratory Testing (ET)

2.1: Have you heard about ET? * (Required, Multiple choice)

Answer categories:
- Yes
- No

2.2: Do you or (some of) your colleagues use ET? * (Required, Multiple choice) Depending on your answer choice, more questions will appear.

Answer categories:
- I use ET. (relevant to SQ 2.3 (a)-2.10)
- I don't use ET but (some of) my colleagues do. (relevant to SQ 2.3 (b))
- I don't use ET and I don't know whether my colleagues do. (relevant to SQ 2.3 (c))
- I don't use ET and none of my colleagues do. (relevant to SQ 2.3 (c))

2.3 (a): Do you use the following testing approaches？ If you do, how frequently do you use the approach with ET? * (Required, Multiple choice)

Answer categories:
- Unit Testing
- Integration Testing
- Regression Testing
- System Testing
- Smoke Testing
- Acceptance Testing
- Security Testing
- Usability Testing
- Performance Testing
- Automated Testing

Levels per answer category:
- No, I don't.
- Yes, I do, but I never use ET to support it.
- Yes, I do, and I use this approach with ET sometimes.
- Yes, I do, and I often use this approach with ET.
- Yes, I do, and I use this approach with ET always.

2.3 (b): Do your colleagues use the following testing approaches？ If they do, how frequently do they use the approach with ET? * (Required, Multiple choice)

Answer categories:
- Unit Testing
- Integration Testing

- Regression Testing
- System Testing
- Smoke Testing
- Acceptance Testing
- Security Testing

Levels per answer category:
- No, they don't.
- I don't know whether they do it.
- Yes, they do, but they never use ET to support it.
- Yes, they do, and they use this approach with ET sometimes.
- Yes, they do, and they often use this approach with ET.
- Yes, they do, and they use this approach with ET always.
- Yes, they do, but I don't know whether they use ET for it.

2.3 (c): How frequently do you use the following testing approachs? * (Required, Multiple choice)

Answer categories:
- Unit Testing
- Integration Testing
- Regression Testing
- System Testing
- Smoke Testing
- Acceptance Testing
- Security Testing
- Usability Testing
- Performance Testing
- Automated Testing

Levels per answer category:
- Never
- Sometimes
- Often
- Always

2.4: In your opinion, which of the following attributes are part of ET? * (Required, Checkboxes)

Answer categories:
- ET has metrics.
- ET has playbooks.
- ET is time-boxed.
- ET is session-based.
- ET has a debriefing meeting.
- ET has systematic coverage tracking.
- ET has a mission statement or a charter.
- ET has a defect log (or defect report/list).
- ET has a test log (recording of what was tested and/or how).
- Other:

- None of above

2.5: When do you typically use ET? * (Required, Multiple choice)

Answer categories:
- I use ET as one of the first test activities.
- I use ET as one of the last test activities.
- I use ET any time during testing.

2.6: Do you use tools that specifically support ET? * (Required, Multiple choice)

Answer categories:
- Yes (Please, list the tool(s)) :
- No

2.7: Do you agree or disagree with the following characterization of ET? * (Required, Multiple choice)

Answer categories:
- ET is easy to use.
- ET supports creativity.
- ET is focused (it's goal-oriented).
- ET makes testing interesting and engaging.
- ET is flexible (can be used in many different test situations).
- ET is efficient (can find defects faster than other methods).
- ET is effective (can find defects which other methods would not).

Levels per answer category:
- Strongly disagree
- Disagree
- Neither agree nor disagree (balanced)
- Agree
- Strongly agree

2.8: Please tell us why you like to use ET. (Up to 3 answers. Please order by priority) (Check boxes, Choose maximum 3 choices)

2.9: Please tell us whether you see any disadvantage with using ET or any situation where you would not recommend using ET. (Up to 3 answers. Please order by priority) (Check boxes, Choose maximum 3 choices)

2.10: Are you planning to change the way how you currently apply ET in order to make it more effective/efficient? If so, please state what you plan to change.

Category 3: Follow-up activities

3.1 Are you willing to provide additional information? * (Required, Multiple choice)
If you agree to share your email address, please check 'Yes', otherwise check 'No'.
Answer categories:
- Yes (relevant to SQ)
- No

3.2 Please leave your email address: * (Required, Only Emails allowed)

B. Data Summary

The "Esi & Fin" and "All" in the following tables indicate the summary results of "Esi & Fin" dataset and "All" dataset.

B.1 Questions on General Background

SQ 1.1: What is your current working role?

Choice	Esi & Fin	Percent 1	All	Percent 2
Tester	26	52.00%	31	50.82%
Test manager	20	40.00%	23	37.70%
Other role cooperating with testers or test managers	4	8.00%	7	11.48%
Answered	50	100.00%	61	100.00%
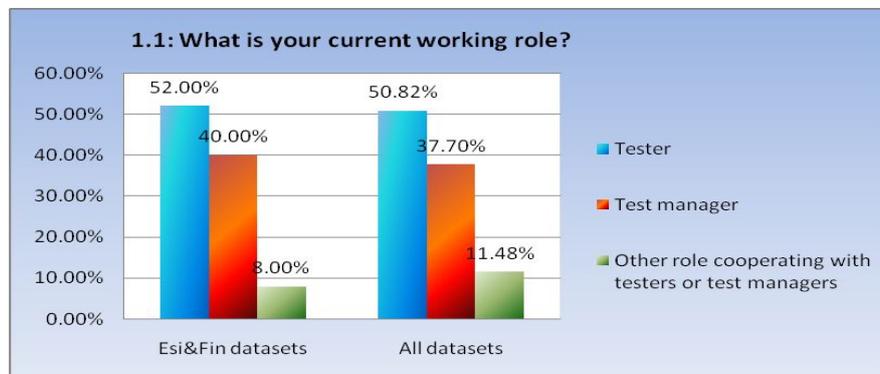
SQ 1.2: How many years have you been working in that role?

Choice	Esi & Fin	Percent 1	All	Percent 2
Less than 2 years	10	20.00%	11	18.03%
2 to 5 years	17	34.00%	20	32.79%
More than 5 years	23	46.00%	30	49.18%
Answered	50	100.00%	61	100.00%
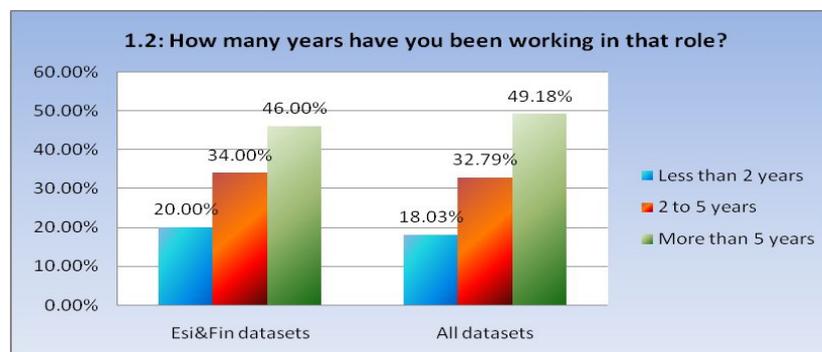
SQ 1.3: Where is your department/unit located?

Choice	Esi & Fin	Percent 1	All	Percent 2
Estonia	27	54.00%	27	44.26%

Finland	23	46.00%	23	37.70%
Other:	0	0.00%	11	18.03%
Answered	50	100.00%	61	100.00%
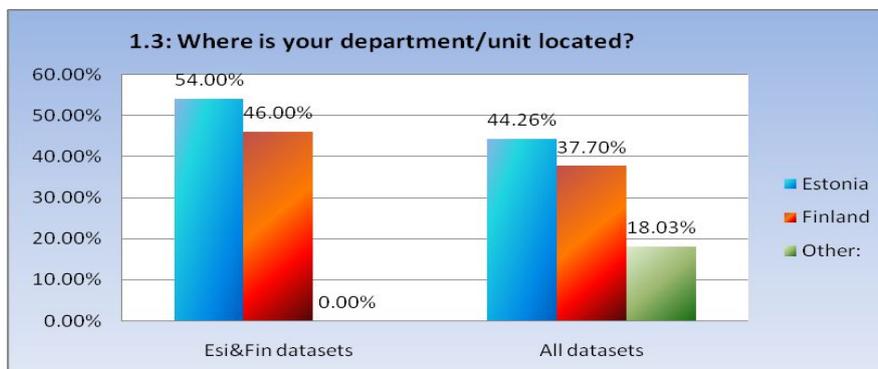
- **11 other respondents from outside Estonia and Finland are:** USA (3), Germany (2), Netherland, Switzerland, India, Belgium, New Zealand, and Sweden.

SQ 1.4: How many employees work in your department/unit?

Choice	Esi & Fin	Percent 1	All	Percent 2
Less than 20	15	30.00%	19	31.15%
20 to 49	16	32.00%	18	29.51%
50 to 99	10	20.00%	12	19.67%
More than 100	9	18.00%	12	19.67%
Answered	50	100.00%	61	100.00%
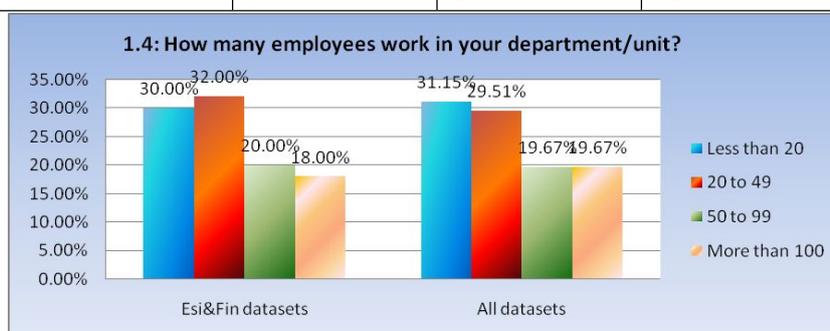
SQ 1.5: How many years has your department/unit existed?

Choice	Esi & Fin	Percent 1	All	Percent 2
Less than 2 years	4	8.00%	7	11.48%
2 to 5 years	19	38.00%	21	34.43%
More than 5 years	27	54.00%	33	54.10%
Answered	50	100.00%	61	100.00%
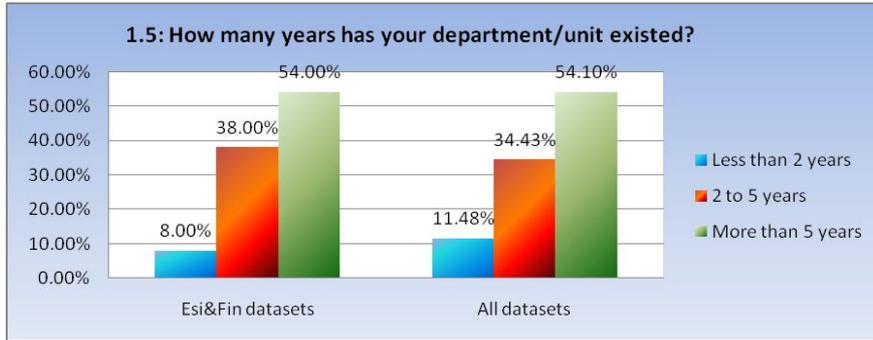
SQ 1.6: Within your company, do you have a separate test (or QA) organization that does testing?

Choice	Esi & Fin	Percent 1	All	Percent 2
Yes	29	58.00%	34	55.74%
No	21	42.00%	27	44.26%
Answered	50	100.00%	61	100.00%
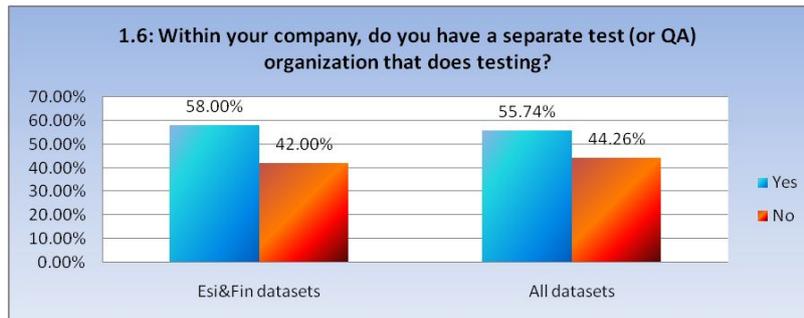
SQ 1.6.1: Are you a member of that separate test (or QA) organization?

Choice	Esi & Fin	Percent 1	All	Percent 2
Yes	24	82.76%	29	85.29%
No	5	17.24%	5	14.71%
Answered	29	58.00%	34	55.74%
Skipped	21	42.00%	27	44.26%
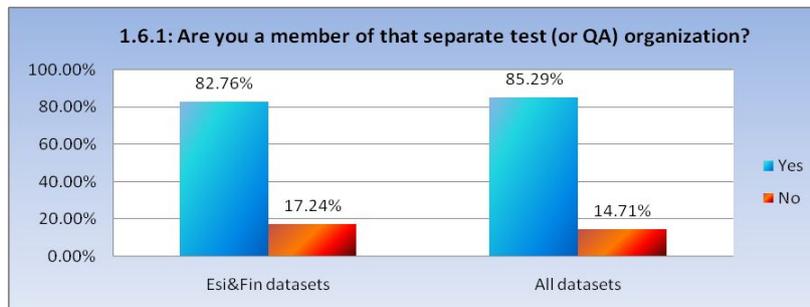
SQ 1.7: What of the following characteristics typically apply to your software?

Choice	Esi & Fin	Percent 1	All	Percent 2

It is usability-critical (e.g., it has a complex GUI which is important for the end user)	41	82.00%	49	80.33%
It has high security demand.	32	64.00%	36	59.02%
It is performance-critical.	36	72.00%	41	67.21%
It is safety-critical.	24	48.00%	26	42.62%
None of above.	2	4.00%	4	6.56%
Answered	50	100.00%	61	100.00%
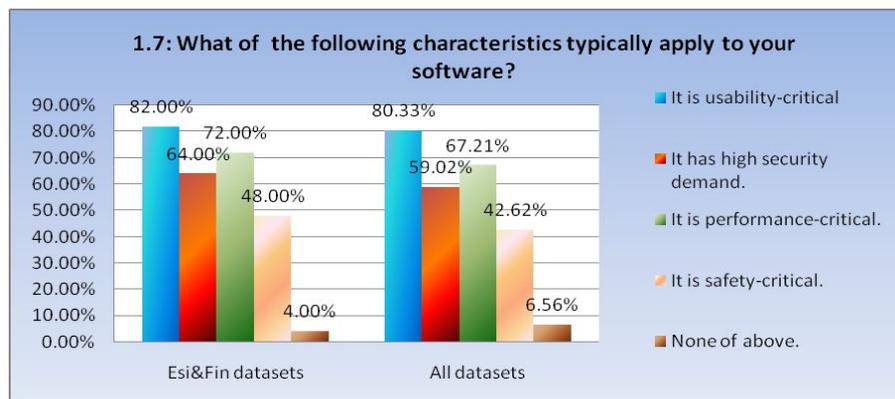
B.2 Questions on Exploratory Testing (ET)

SQ 2.1: Have you heard about ET?

Choice	Esi & Fin	Percent 1	All	Percent 2
Yes	50	100.00%	61	100.00%
No	0	0.00%	0	0.00%
Answered	50	100.00%	61	100.00%

SQ 2.2: Do you or (some of) your colleagues use ET?

Choice	Esi & Fin	Percent 1	All	Percent 2
I use ET. (relevant to Question 2.3 (a)-2.10)	44	88.00%	54	88.52%
I don't use ET but (some of) my colleagues do. (relevant to Question 2.3 (b))	1	2.00%	2	3.28%
I don't use ET and I don't know whether my colleagues do. (relevant to Question 2.3 (c))	5	10.00%	5	8.20%
I don't use ET and none of my colleagues do. (relevant to Question 2.3 (c))	0	0.00%	0	0.00%
Answered	50	100.00%	61	100.00%
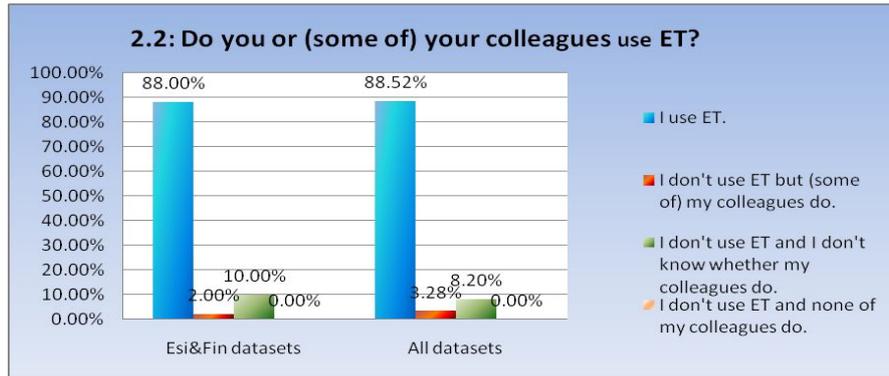
SQ 2.3 (a): Do you use the following testing approaches？If you do, how frequently do you use the approach with ET?

	No, I don't.	Yes, I do, but I never use ET to support it.	Yes, I do, and I use this approach with ET sometimes.	Yes, I do, and I often use this approach with ET.	Yes, I do, and I use this approach with ET always.	Total
Unit Testing	40	2	2	4	6	54
Integration Testing	10	3	12	14	15	54
Regression Testing	2	2	8	21	21	54
System Testing	2	1	7	16	28	54
Smoke Testing	6	6	9	8	25	54
Acceptance Testing	9	3	12	12	18	54
Security Testing	19	4	10	7	14	54
Usability Testing	8	2	5	16	23	54
Performance Testing	16	12	10	8	8	54
Automated Testing	13	21	6	6	8	54
	125	56	81	112	166	

SQ 2.3 (b): Do your colleagues use the following testing approaches？If they do, how frequently do they use the approach with ET?

	No, they don't.	I don't know whether they do it.	Yes, they do, but they never use ET to support it.	Yes, they do, and they use this approach with ET sometimes.	Yes, they do, and they often use this approach with ET.	Yes, they do, and they use this approach with ET always.	Yes, they do, but I don't know whether they use ET for it.	Total
Unit Testing	0	1	1	0	0	0	0	2
Integration	0	0	0	1	1	0	0	2

Testing								
Regression Testing	0	0	0	1	1	0	0	2
System Testing	0	0	0	1	1	0	0	2
Smoke Testing	0	0	0	1	1	0	0	2
Acceptance Testing	0	0	0	1	1	0	0	2
Security Testing	1	0	0	0	1	0	0	2
Usability Testing	0	0	0	1	0	1	0	2
Performance Testing	0	0	0	1	1	0	0	2
Automated Testing	0	0	0	1	1	0	0	2

SQ 2.3 (c): How frequently do you use the following testing approaches?

	Never	Sometimes	Often	Always	Total
Unit Testing	1	2	2	0	5
Integration Testing	0	4	1	0	5
Regression Testing	0	3	2	0	5
System Testing	0	1	3	1	5
Smoke Testing	1	3	1	0	5
Acceptance Testing	1	2	2	0	5
Security Testing	3	1	1	0	5
Usability Testing	3	1	1	0	5
Performance Testing	1	3	1	0	5
Automated Testing	1	2	1	1	5
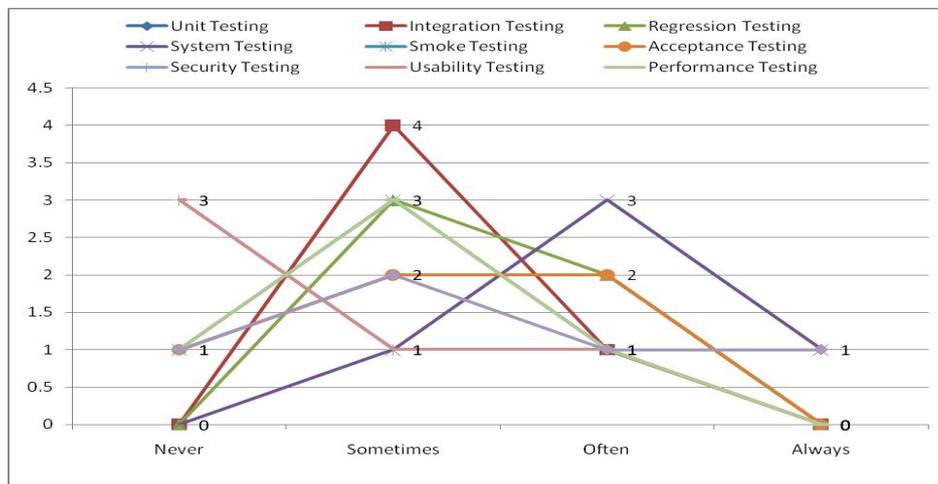
SQ 2.4: In your opinion, which of the following attributes are parts of ET?

Choice	Esi & Fin	Percent 1	All	Percent 2
ET has metrics.	10	22.73%	14	25.93%
ET has playbooks.	7	15.91%	12	22.22%
ET is time-boxed.	21	47.73%	28	51.85%
ET is session-based.	24	54.55%	33	61.11%
ET has a debriefing meeting.	15	34.09%	20	37.04%
ET has systematic coverage tracking.	10	22.73%	14	25.93%
ET has a mission statement or a charter.	23	52.27%	31	57.41%
ET has a defect log (or defect report/list).	35	79.55%	42	77.78%
ET has a test log (recording of what was tested and/or how).	31	70.45%	39	72.22%
Other:	9	20.45%	12	22.22%
None of above	2	4.55%	2	3.70%
Answered	44	88.00%	54	88.52%
Skipped	6	12.00%	7	11.48%
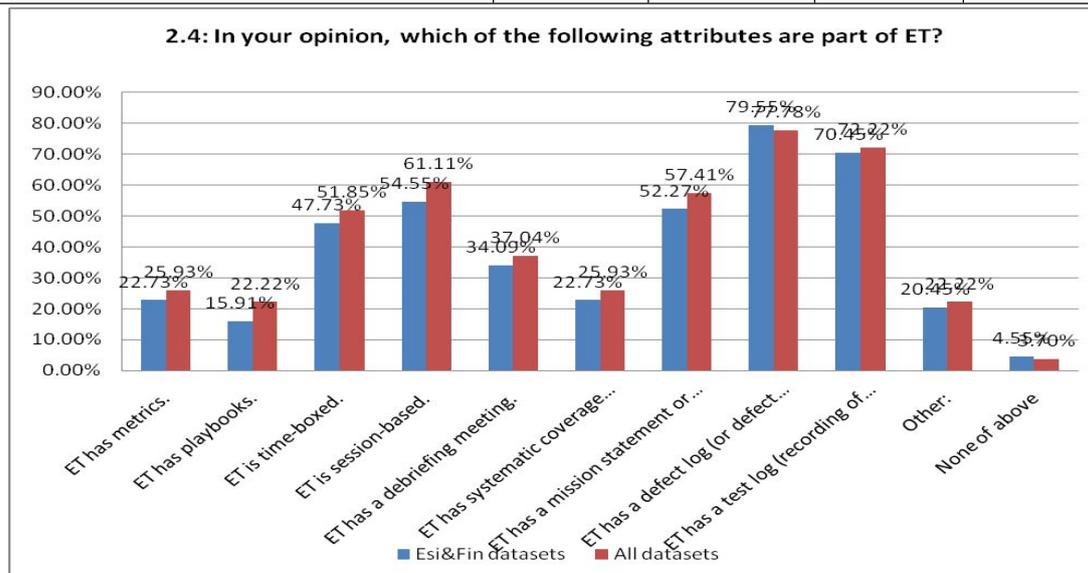
- **Other answers:**
  - ❖ Extra testing in addition to planned testing.
  - ❖ Don't mix ET with Session Based Testing. The plainest definition of exploratory testing is test design and test execution at the same time. ET is an approach, not another testing technique.
  - ❖ The statements above are about SBTM not ET. ET can be done without all of the above or with some of the above.
  - ❖ ET utilizes people and tools
  - ❖ ET can have all those things, but not necessariliy always together
  - ❖ Based on oracles, skills, ideas etc
  - ❖ Catches bug which other testing types misses.
  - ❖ Sapience

- ❖ ET is about simulatinous exploration, observing, planning, experimentation and communicating your findings. All these running in a loop followed by each other. Purpose is to test by exploring.
- ❖ Tesing with up to date requirements
- ❖ All of the options could be used with ET but they seem to be attributes of STBM, not ET in itself
- ❖ ET can have anything you want. It's an approach to testing.

SQ 2.5: When do you typically use ET?

Choice	Esi & Fin	Percent 1	All	Percent 2
I use ET as one of the first test activities.	7	15.91%	10	18.52%
I use ET as one of the last test activities.	5	11.36%	5	9.26%
I use ET any time during testing.	32	72.73%	39	72.22%
Answered	44	88.00%	54	88.52%
Skipped	6	12.00%	7	11.48%
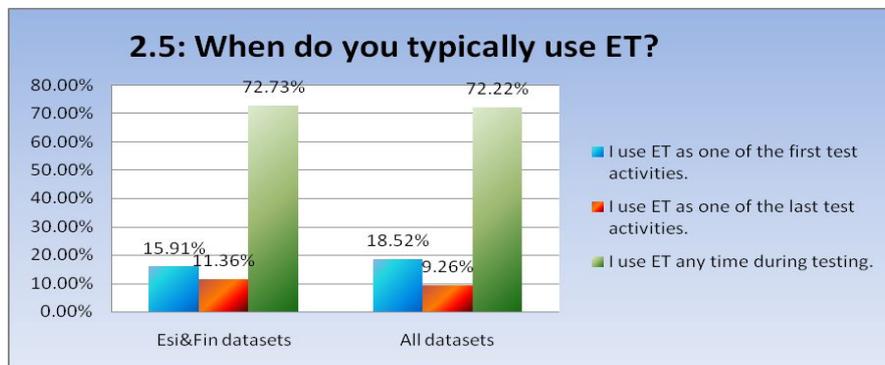
SQ 2.6: Do you use tools that specifically support ET?

Choice	Esi & Fin	Percent 1	All	Percent 2
Yes (Please, list the tool(s)) :	11	25.00%	15	27.78%
No	33	75.00%	39	72.22%
Answered	44	88.00%	54	88.52%
Skipped	6	12.00%	7	11.48%
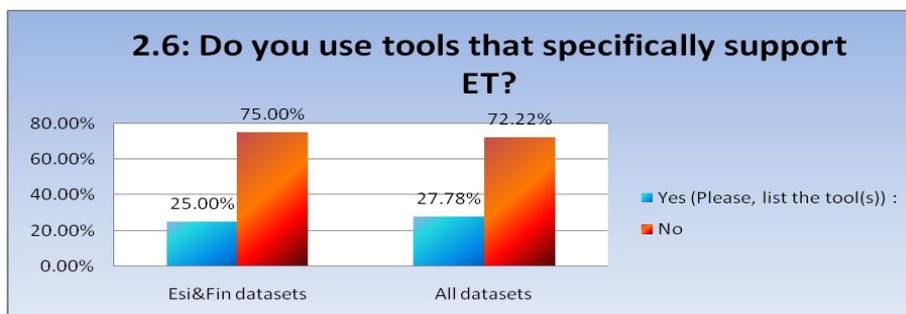
- ● **Tools that specifically support ET**
  - ❖ ET testing charter and ET testing result template
  - ❖ Any mindmap software.

- ❖ Excel, custom python scripts, custom perl scripts, BB Flashback, rapid reporter, mindmaps, Xmind
- ❖ perlclip, IETester
- ❖ note taking tools like OneNote or Evernote, mindmapping tools like XMind, tools like allpairs and counterstring
- ❖ Custom made Session management tool
- ❖ Rapid Reporter, Mind Maps, Evernote, Screen Recorder
- ❖ Jira Test Sessions
- ❖ XMind mind mapping tool
- ❖ Vim-Editor
- ❖ Check lists, own experience, literature
- ❖ Rapid Reporter is a good example if the question is about what tools are made to support an exploratory tester.
- ❖ Microsoft Test Manager
- ❖ Rapid Reporter
- ❖ XMind, Rapid Reporter, qTrace, pen & paper, PostIts, brains, etc.

SQ 2.7: Do you agree or disagree with the following characterization of ET?
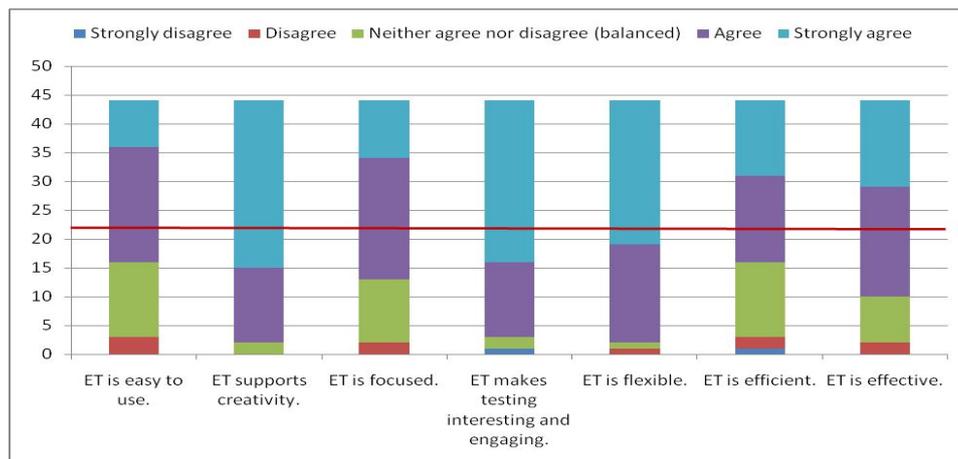
The median, 44 times 50%, is 22, so we try to find out the option of each item in cumulative frequency is equal to 22[5]. Form the graph below, we can see respondents choose "Agree" or "Strongly agree" for each option we gave. Especially, respondents strongly agree "ET supports creativity", "ET makes testing interesting and engaging" and "ET is flexible".

a) Esi & Fin dataset

	Strongly disagree	Disagree	Neither agree nor disagree (balanced)	Agree	Strongly agree	Total
ET is easy to use.	1	3	15	22	13	54
ET supports creativity.	0	0	2	15	37	54
ET is focused (it's goal-oriented).	0	3	11	25	15	54
ET makes testing interesting	1	0	2	15	36	54

and engaging.						
ET is flexible (can be used in many different test situations).	0	1	2	20	31	54
ET is efficient (can find defects faster than other methods).	1	3	14	17	19	54
ET is effective (can find defects which other methods would not).	1	3	10	21	19	54

	-2	-1	0	1	2	Total
ET supports creativity.	0	0	0	13	59	72
Interesting and engaging.	-2	0	0	13	56	67
ET is flexible	0	-1	0	17	50	66
ET is effective	0	-2	0	19	30	47
ET is focused	0	-2	0	21	20	39
ET is efficient	-2	-2	0	15	26	37
ET is easy to use.	0	-3	0	20	16	33

b) All dataset

	Strongly disagree	Disagree	Neither agree nor disagree (balanced)	Agree	Strongly agree	Total
ET is easy to use.	0	3	13	20	8	44
ET supports creativity.	0	0	2	13	29	44
ET is focused (it's goal-oriented).	0	2	11	21	10	44
ET makes testing interesting and engaging.	1	0	2	13	28	44
ET is flexible (can be used in many different test situations).	0	1	1	17	25	44
ET is efficient (can find defects faster than other methods).	1	2	13	15	13	44
ET is effective (can find defects which other methods would not).	0	2	8	19	15	44

	-2	-1	0	1	2	Total
ET supports creativity.	0	0	0	15	74	89
Interesting and engaging.	-2	0	0	15	72	85
ET is flexible	0	-1	0	20	62	81

ET is effective	-2	-3	0	21	38	54
ET is focused	0	-3	0	25	30	52
ET is efficient	-2	-3	0	17	38	50
ET is easy to use.	-2	-3	0	22	26	43
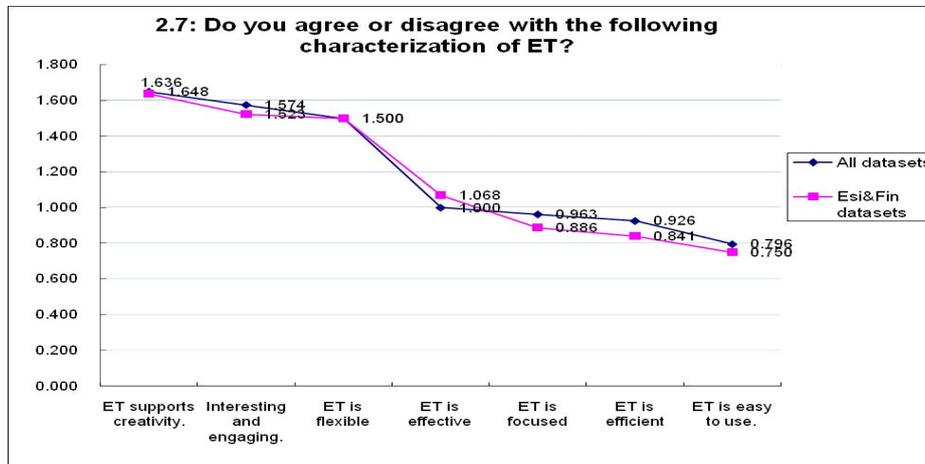
We can see from above graph two datasets have quite similar trend.

SQ 2.8: Please tell us why you like to use ET. (Up to 3 answers. Please order by priority)

Advantage 1	
Finds bugs more effectively	effective
Find new bugs	effective
Adapte to the changes of the target software	flexible
In exploratory testing the tester has to ask why this part needs to be tested and what is important in it.	emphasizes tester
Fast,	efficient
Can find defects faster	efficient
Discover unexpected bugs due to the very nature of exploratory testing	effective
Thinking outside the box	supports creativity
It keeps me intellectually engaged	supports creativity
Additional code coverage	flexible
Supports learning	supports learning
ET engages the tester's mind and requires critical thinking which helps to find gaps and potential issues in the product.	emphasizes tester
quicker, when focused correctly	efficient
When I apply (!) ET to my testing, I use my brains instead of a script	supports creativity
ET is a good complementary method to catch defects that other methods may have missed.	effective
Efficient	efficient
ET can achieve a large coverage with minimal time required	efficient
To try to break the system after everything is tested according to software test plans	effective

Defects find quickly	efficient
Method is practical - no unnecessary waste like nice words in test plans for upper management	easy
when you're lacking time to prepare detailed scripted test cases	easy
Helps to have more time doing actual testing and not writing very long plans and documentation	time saving
It saves me from documenting complex test activities on a detailed level (test steps, etc)	time saving
Exploring a new feature is fastest way to start finding bugs	efficient
This allows both extended scenario based and out of the box testing, which adds variety and is interesting.	interesting
Best way to regression test if only minor changes have been made to any specific feature and there is no need to run the full spec.	focused
Time is not wasted on writing test cases	time saving
ET supports creativity	supports creativity
Finds bugs	effective
An extra mile to find that unusual bug that might not be found in the normal test-script	effective
It's easy to implement	easy
More real testing and less formal procedures	efficient
I learn and think during testing, so it kind of fits naturally.	supports learning
helps to understand better how system works	supports learning
Fun	interesting
Supports creative thinking	supports creativity
Its one of the most efficient ways to gather information regarding how particular software actually works	efficient
ET makes testing execution more attractive and versatile for the tester	supports creativity
I can produce valuable results immediately, not after reading through a specification or going through various test cases or something such	time saving
You already have above some, but mainly because testing can't be done without exploratory activities. It's just not possible.	essential
Finding important defects faster	efficient
Gives the tester personal freedom and responsibility to optimise his/her work	flexible

Advantage 2	
streches brains, makes testing more funny	interesting
Scalable	supports creativity
Motivate tester thinking continiously	supports creativity
Tester needs to be able to justify what has been done and why it is enough.	supports creativity
Productive	effective

Can save time (planning, test execution)	time saving
Using pre-written test cases would not be efficient, as our iterations are small and subject to change	efficient
Users do the craziest things	supports creativity
Fastest feedback of relevant information	efficient
Since we don't have written specs, ET works well for finding information even if you have limited information available about the product.	effective
improves test assets	efficient
When I apply (!) ET to my testing and use a script, I can utilize my critical thinking and creativity	supports creativity
Helps to learn product	supports learning
Creativity in use	supports creativity
Allows to go beyond the predefined scripted test cases (allows creativity in the scope definition)	supports creativity
Is more flexible than scripted tests - allows more room for exploring other parts of the application	flexible
It is easiest way to simulate real user behaviour	easy
This allows creative thinking and suits every level of skills.	supports creativity
Best way to quickly find defects with new features	efficient
New test ideas that pop up during testing can be instantly executed	efficient
ET is effective (can find defects which other methods would not)	effective
helps to teach/mentor testers	supports learning
Tester is urged to think out of the box	supports creativity
Independent	Independent
Is not boring, supports creativity. Boring things should be automated.	interesting & supports creativity
makes testing more interesting	interesting
Makes brain work	supports creativity
Is an efficient method that is easy to approach	efficient & easy
It encourages to vary your tests as much as possible as otherwise it is not exploration anymore	flexible
Testing with the latest and up to date requirements in the fast changing enviroments	flexible
I *love* the exploring. One of my strongest points is my imagination and, with ET, I can utilize it to its maximum.	supports creativity
I can learn about the product while using it	supports learning
Tester's skill becomes central focus of testing rather than tool	emphasizes tester

Advantage 3	
allow tester to decide, feel of (limited) freedom instead of prewritten exact steps	supports creativity
Flexibility	flexible

Exploratory testing emphasizes the responsibility of the tester and also provokes tester to find interesting and important thing	emphasizes tester
Save Time	time saving
Create logs (by using recording tool)	create logs
ET helps me gain a deeper understanding of the product than simply running test cases would	supports learning
User data is never neat and tidsy	clear data
Highly customisable	focused
teaches a lot to testers and developers	supports learning
Creativity	supports creativity
I get less bored during testing, running the same set of scripted tests should be automated whenever possible	interesting
ET allows a fresh view on known software and capitalizes the fresh views of a novice in usage of this software too.	supports creativity
Best way to verification test fixed defects and possibly affected areas	effective
Do not set limits on ideas	supports creativity
ET makes testing interesting and engaging	interesting
keeps up the motivation	interesting
Catches few commonly missed bugs	effective
helps to find bugs that would not be found otherwise	effective
Efficient	efficient
Applicable to many different testing situations	flexible
In order to provide value from testing you must communicate. In ET you need to keep a testlog to effectively communicate of your finding. Thus this leads to more effective testing too if logs are reviewed and thought about later.	effective
Testing has some randomness to find some rare defects	effective
It allows me to be creative	supports creativity

SQ 2.9: Please tell us whether you see any disadvantage with using ET or any situation where you would not recommend using ET. (Up to 3 answers. Please order by priority)

Disadvantage 1	
some tests by authorities are defined in every detail, they must be run exactly in that way.	inflexible
Somtimes Focus might shifted	no focus
Provide estimates and timelines	time limit
Difficult for reporting to managment as there are no hard matrics (pass/fail rates, etc.)	hard to report
Harder to compare results to previous ones	hard to compare results
ET might be difficult when the application under test is very	not good for complicated

complex and the tester is not familair with the application.	project
Potential for incomplete testing if used exclusivly	not all-inclusive
People using it as a buzzword and not underestanding it.	too popular
stakeholders don't appreciate	stakeholders don't appreciate
Do not use it if you do not have a good system to record your progress or estimate coverage.	hard to record
Do not suit for everyone (new testers, un-motivated employee etc).	high requirement for tester
ET may not find more technical or in depth problems	not good for complicated project
I would never recommend ET for technical projects	not good for complicated project
When there are complicated calculation etc. rules to be tested	not good for complicated project
To be effective, requires quite a lot knowledge about substance area and method itself	high requirement for tester
In case of inexperienced and not very creative emplyees it's difficult to make them think in a systematic and \"out of the box\" way	high requirement for tester
Harder to repeat the same set of tests because every session is unique	unrepeatable
ET does not save you from doing test planning. Poorly planned ET is can be worse than poorly planned \"scripted testing\".	time consuming
You need to have a proper out of the box mindset and creativity, simply covering same things as in scripted testing does not work	high requirement for tester
ET can never be used as a standalone test technique, cannot guarantee expected quality level, must be combined with spec-based manual or automated tests	not all-inclusive
Needs right mindset and skills to execute properly	high requirement for tester
when you have only novice people in testing team and no senior tester	high requirement for tester
The chance of spending time without actually finding anything	inefficient
During Automation.. You can't script to explore random	inflexible
More suitable for experienced testers; juniors tend to cover less critical areas when using ET	high requirement for tester
I would not use any kind of manual testing when automated checking does the job. For example smoke tests.	inflexible
can be more time consuming sometimes	time consuming
Organizations that have strict testing policies and fixed processes	inflexible
If the person testing is not trained on using exploration and note-taking for testing. Then the results from that testing might even be poorer than static execution of detailed test steps (altought	high requirement for tester

static execution of test steps is impossile for humans and there is always \"free exploration\" -element in evererything we do)	
Requires a lot of experience from the tester	high requirement for tester
It is widely misunderstood (based on the \"ET is/has\" options, even in this survey). When done properly, ET is not unstructured, it is not chaotic and it does not forego reporting.	confusing
I can't imagine such a situation because it's not possible. Testing doesn't exist without exploratory component.	unnecessary
difficult without a GUI - for example in web services testing	inflexible

Disadvantage 2	
too loose goals may leave large uncovered areas	no focus
Might hit some issue which is not relevant	uncontrollable
Show coverage and coverage strategy	hard to record
Management often doesn't value the time spent in ET as opposed to formal scripted testing	stakeholders don't appreciate
Requires a tester with good communication skills.	high requirement for tester
takes time to teach testers the right mindset to execute ET	high requirement for tester
Hard to control & manage	uncontrollable
No record of detailer test report	hard to record and report
Mored difficult to get reports about which tests were run	hard to report
Sometimes difficult to repro situations.	hard to report
For ET tester has to know the software well	high requirement for tester
Sometimes difficult to track coverage	hard to record
when you have asked to switch off your brain	energy consuming
The tendency to do the test based on assumtion	inaccurate
No detailed test plans/reports but some clients ask for detailed test documentation	hard to record and report
can be less effective if used in wrong place etc	ineffective
Testers with very limited experience in testing	high requirement for tester
Effective exploration is surprisingly hard to conduct effectively as humans have tendency to start following safe and already explored patterns.	ineffective
First point may easily result in people misusing the term and, as a side-effect, cause people to dislike ET because they've only seen it done badly.	too popular

Disadvantage 3	
tester may get deep in some areas and there is not enough time for others demanded.	time limit

| Test data strategy and coverage is not inherent | inflexible |
| Depends more on the current mood of the tester (I personally tend to be less creative with my testing in the evening period) | high requirement for tester |

SQ 2.10: Are you planning to change the way how you currently apply ET in order to make it more effective/efficient? If so, please state what you plan to change.

Plan to change	
I'm thinking of how to more easily document where I have been during the ET session. I would prefer using tester pair, one tests, another just writes down.	Create a record for ET
Of course, the way of doing ET may vary according to the software and projects.	Choose the way of doing ET according to the project
Planning to implment Matrics, and also allocate and stick to time frame for ET	Create a record for ET
Understand more about ET and improve better reporting of ET results so they can be used as effective means to determine the overall quality of the product.	Improve report
No plans at the moment; however, we usually adjust our approach to changes in the project.	no
Nope, not planning to change the approach to testing.	no
I am planning to experiment with SBTM and TBTM to find a good balance for using ET. I also want testers to work on their individual skills to be better explorers.	Use SBTM and TBTM together
I am planning to make my testing more credible by applying more recordings of my doings.	Create a record for ET
I want to create or find a better system for exploratory test session report.	Find a better reporting system for ET
No	no
No.	no
I plan to make it more structured by applying either SBTM or some other way to save session logs and then have a log to view when needed. Also might try to apply some metrics to the session logs/outcomes. I plan also to conduct ET activities mostly during the morning and try to come up with additional charters/tours for my product.	Improve report and do ET in the morning
Yes, just use ET more often.	Use ET more often
Use continuous improvement all the time to improve testing	Improve testing all the time
At this moment there's little room with ET with so little time frame, so then answer is no.	no
More risk-based testing.	More risk-based testing.
I could always use more automation and better tools, something like a code coverage report for ET sessions or faster data entry with Webdriver. No	Use a good tool

specific plans yet, since I just started on a new project.	
I'm not planning to change anything at the moment.	no
Ones needs to continually study their craft. For me (area of soft. testing) this relates to need to study things like: Note-taking and observation or study ways to create innovative yet rigid experiments.Testing is about experimentation, observation and communication. These are the main areas testers need to study.	Study more and have more experience
I am planning to change a lot of the reporting my department is doing in order to a) report more interesting things, b) decrease time used for reporting, and c) make it clear for testers that they own their work so they need to come up with proposals.	Make report more effective and efficient

B.3 Follow-up Activities

SQ 3.1: Are you willing to provide additional information?

Choice	Count	Percent
Yes	26	42.62%
No	35	57.38%
Answered	61	100.00%

C. SPSS Analysis Results

C.1 "All" Dataset Group

***RQ 2.1** Do testers use ET more frequently than other roles?*

Role * Reply Crosstabulation

Count

		Reply		Total
		yes	no	
Role	Tester	29	2	31
	Tester Manager	20	3	23
	Other roles	5	2	7
Total		54	7	61

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	2.840[a]	2	.242	.244		
Likelihood Ratio	2.455	2	.293	.433		
Fisher's Exact Test	2.949			.244		
Linear-by-Linear Association	2.572[b]	1	.109	.143	.098	.065
N of Valid Cases	61					

a. 3 cells (50.0%) have expected count less than 5. The minimum expected count is .80.
b. The standardized statistic is 1.604.

P=0.244 > 0.05, accept H$_0$, which means there are no significant differences in different roles that use ET.

***RQ 2.2** Do software engineers with more working experience use ET more frequently?*

Year * Reply Crosstabulation

Count

		Reply		Total
		Yes	No	
Year	< 2	8	3	11
	2-5	17	3	20
	> 5	29	1	30
Total		54	7	61

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	4.905[a]	2	.086	.079		
Likelihood Ratio	4.906	2	.086	.139		
Fisher's Exact Test	4.910			.063		
Linear-by-Linear Association	4.823[b]	1	.028	.034	.030	.021
N of Valid Cases	61					

a. 3 cells (50.0%) have expected count less than 5. The minimum expected count is 1.26.
b. The standardized statistic is -2.196.

P=0.063 >0.05, so accept H0, reject H1, which means there is no relevance between working experience and using ET.

Even though the results from two datasets are different, the P value from "All" dataset is quite close to 0.05.

RQ 2.3 Does the location (Estonia or Finland) have an effect on applying ET?

Location * Reply Crosstabulation

Count

		Reply		Total
		Yes	No	
Location	Estonia	25	2	27
	Finland	19	4	23
	Others	10	1	11
Total		54	7	61

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	1.294[a]	2	.524	.586		
Likelihood Ratio	1.259	2	.533	.586		
Fisher's Exact Test	1.272			.664		
Linear-by-Linear Association	.200[b]	1	.655	.791	.420	.185
N of Valid Cases	61					

a. 3 cells (50.0%) have expected count less than 5. The minimum expected count is 1.26.
b. The standardized statistic is .447.

P=0.664 > 0.05, accept H0. There is no significant difference in different locations of the department that applies ET.

RQ 2.4 Do software engineers in larger companies use ET more often than in smaller companies?

Company size * Reply Crosstabulation

Count

		Reply		Total
		Yes	No	
Number of employees	< 20	18	1	19
	20-49	16	2	18
	50-99	9	3	12
	> 100	11	1	12
Total		54	7	61

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	3.001[a]	3	.391	.440		
Likelihood Ratio	2.700	3	.440	.570		
Fisher's Exact Test	2.734			.421		
Linear-by-Linear Association	.547[b]	1	.459	.479	.287	.106
N of Valid Cases	61					

a. 4 cells (50.0%) have expected count less than 5. The minimum expected count is 1.38.
b. The standardized statistic is .740.

P=0.421 > 0.05, accept H0. There is no significant difference in different sizes of the company that applies ET.

RQ 2.5 *Is ET used more frequently in older companies?*

Company history * Reply Crosstabulation

Count

		Reply		Total
		Yes	No	
Company history (years)	< 2	6	1	7
	2-5	18	3	21
	> 5	30	3	33
Total		54	7	61

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	.402[a]	2	.818	.862		
Likelihood Ratio	.401	2	.818	.862		
Fisher's Exact Test	.823			.737		
Linear-by-Linear Association	.324[b]	1	.569	.776	.376	.182
N of Valid Cases	61					

a. 3 cells (50.0%) have expected count less than 5. The minimum expected count is .80.
b. The standardized statistic is -.569.

P=0.737 > 0.05, accept H0. There is no significant difference longer or shorter age of the company that applies ET.

RQ 2.6 *Do software engineers in separate test organizations use ET more frequently?*

Separate organization * Reply Crosstabulation

Count

		Reply		Total
		Yes	No	
Separate organization	Have	27	2	29
	Don't have	3	2	5
Total		30	4	34

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	4.502[a]	1	.034	.094	.094	
Continuity Correction[b]	1.878	1	.171			
Likelihood Ratio	3.345	1	.067	.094	.094	
Fisher's Exact Test				.094	.094	
Linear-by-Linear Association	4.370[c]	1	.037	.094	.094	.088
N of Valid Cases	34					

a. 3 cells (75.0%) have expected count less than 5. The minimum expected count is .59.
b. Computed only for a 2x2 table
c. The standardized statistic is 2.090.

P= 0.094 > 0.05, accept H0, There is no significant difference whether there is a separate test organization or not in the company that applies ET.

C.2 Other Results for RQ2.7

Role * Reply Crosstabulation

Count

		Reply		Total
		Estonia	Finland	
Role	Tester	14	10	24
	Tester manager	11	7	18
	Other roles	0	2	2
Total		25	19	44

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	2.789[a]	2	.248	.352		
Likelihood Ratio	3.518	2	.172	.322		
Fisher's Exact Test	2.390			.352		
Linear-by-Linear Association	.597[b]	1	.440	.608	.302	.151
N of Valid Cases	44					

a. 2 cells (33.3%) have expected count less than 5. The minimum expected count is .86.
b. The standardized statistic is .773.

Year * Reply Crosstabulation

Count

		Reply		Total
		Estonia	Finland	
Year	< 2	4	3	7
	2-5	10	4	14
	> 5	11	12	23
Total		25	19	44

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	1.976[a]	2	.372	.385		
Likelihood Ratio	2.023	2	.364	.385		
Fisher's Exact Test	1.983			.385		
Linear-by-Linear Association	.720[b]	1	.396	.426	.261	.114
N of Valid Cases	44					

a. 2 cells (33.3%) have expected count less than 5. The minimum expected count is 3.02.
b. The standardized statistic is .849.

Separate organization * Reply Crosstabulation

Count

		Reply		Total
		Estonia	Finland	
Separate organization	Yes	11	11	22
	No	1	2	3
Total		12	13	25

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	.294[a]	1	.588	1.000	.531	
Continuity Correction[b]	.000	1	1.000			
Likelihood Ratio	.300	1	.584	1.000	.531	
Fisher's Exact Test				1.000	.531	
Linear-by-Linear Association	.282[c]	1	.595	1.000	.531	.407
N of Valid Cases	25					

a. 2 cells (50.0%) have expected count less than 5. The minimum expected count is 1.44.
b. Computed only for a 2x2 table
c. The standardized statistic is .531.

Company history * Reply Crosstabulation

Count

		Reply		Total
		Estonia	Finland	
History (years)	< 2	2	1	3
	2-5	9	7	16
	> 5	14	11	25
Total		25	19	44

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	.128[a]	2	.938	1.000		
Likelihood Ratio	.131	2	.937	1.000		
Fisher's Exact Test	.269			1.000		
Linear-by-Linear Association	.059[b]	1	.809	1.000	.503	.186
N of Valid Cases	44					

a. 2 cells (33.3%) have expected count less than 5. The minimum expected count is 1.30.
b. The standardized statistic is .242.