Trento 2014 Security Engineering Course: Experiment Description

April 17, 2014

1 Research method

This section describes the design of the performed experiment, following the guidelines by Wohlin et al. [8].

1.1 Research Questions

The *goal* of the experiment was to compare visual and textual methods for security risk assessment with respect to how successful they are in identifying threats and security controls. For this purpose we have adopted as dependent variables the success constructs defined in the Method Evaluation Model (MEM) proposed by Moody [6]: *effectiveness*, *perceived easy of use*, *perceived usefulness*, and *intention to use*. Therefore, we have specified the following research questions that match the constructs of the MEM:

RQ1 *Is the effectiveness of the methods significantly different between the two types of methods?*

RQ2 *Is the effectiveness of the methods significantly different between the two facets?*

RQ3 *Is the participants' overall perception of the method significantly different between the two type of methods?*

RQ4 *Is the participants' perceived usefulness of the method significantly different between the two type of methods?*

Variable	**Scale**	**Means**	**Distribution**
Gender	Sex		79% were male; 21%were female
Age	Years	25.72	48% were 21-24 years; 41% were 25-29; 10% were 30-40
Education Length	—"—	4.28	66% had <5 years; 17% had 5 years; 17% had >5 years
Work Experience	—"—	2.46	31% had no experience; 31% had < 2 years; 28% had 3-5 years; 10% had >6 years
Level of Expertise in Security Technology	1(Novice)-5(Expert)	2.31	28% novices; 28% beginners; 10% competent users; 31% proficient users; 3% experts
Level of Expertise in Security Regulation and Standards	—"—	1.86	45% novices; 17% beginners; 7% competent users; 31% proficient users
Level of Expertise in Privacy Technology	—"—	2.10	31% novices; 34% beginners; 28% competent users; 7% proficient users
Level of Expertise in Privacy Regulation	—"—	1.90	48% novices; 24% beginners; 7% competent users; 21% proficient users
Level of Expertise in RE	—"—	2.31	24% novices; 34% beginners; 14% competent users; 28% proficient users

Table 1: Demographic Statistics

RQ5 *Is the participants' perceived ease of use of the method significantly different between the two type of methods?*

RQ6 *Is the participants' intention to use the method significantly different between the two type of methods?*

We have translated research questions $RQ1 - RQ6$ into a list of null hypotheses to be statistically tested. We do not list them here due to the lack of space. To answer $RQ1$ and $RQ2$ we have measured methods' *actual effectiveness* by counting the number of threats and security controls identified with each method application and we asked an external security expert to assess their quality. Research questions $RQ3$-$RQ6$ have been answered by administering to the participants a post-task questionnaire inspired to the Method Evaluation Model (MEM) [6] after they have completed each of the method applications. To gain a better understanding *why there is a difference in methods effectiveness and perception* we have conducted individual interview with participants.

1.2 Methods Selection

As in our previous experiment [4], we have chosen as instance of the visual method CORAS [5] because is the only visual method for security risk assessment. CORAS is a method designed at SINTEF, a research institution in Norway which is used to provide security risk assessment consulting services. It consists of three tightly integrated parts, namely, a method for risk analysis, a language for risk modeling, and a tool to support the risk analysis process. The risk analysis in CORAS is a structured and systematic process which uses diagrams (see Figure **??**) to document the result of the execution of each step. The steps are based on the international standard ISO 31000 [3] for risk management: context establishment, risk analysis (that identifies assets, unwanted incidents, threats and vulnerabilities), and risk treatments. Instead, we have replaced SREP, the instance of textual method used in the original experiment, with SecRAM [2], an industrial method by EUROCONTROL used to conduct security risk assessment in the air traffic management domain (ATM). SecRAM supports the security risk management process for a project initiated by an air navigation service provider, or ATM project, system or facility. SecRAM provides a systematic approach to conduct security risk assessment which consists of five main steps: defining the scope of the system, assessing the impact of a successful attack, estimating the likelihood of a successful attack, assessing the security risk to the organization or project, and defining and agreeing a set of management options. As shown in Figure **??**) tables are used to represent the results of each step's execution.

1.3 Domain Selection

We selected the Smart Grid application scenario for our experiment as we had already used in the previous experiment so that we could compare the results from the two experiments. The Smart Grid is an electricity network that uses information and communication technologies to optimize the distribution and transmission of electricity from supply points to end-consumers. The application scenario focused on the gathering of metering information from the smart meters located in private households and its communication to the electricity supplier for billing purposes.

1.4 Demographics

The participants of the experiment were recruited among MSc students enrolled in the Security Engineering course at the University of Trento. Table 1 presents descriptive statistics about the participants. Most of the participants (69%) reported that they had at least 2 years of working experience

Table 2: Original experiment and replication settings

	Original	**Replication**
Subject Type	28 MSc students	29 MSc students
Subject Unit	16 groups of 1-2 students	29 Groups of 1 student
Subject Environment	Security Engineering course	Security Engineering course
Experiment Task	Identify threats & controls	Identify threats & controls
Time to complete the task	4 sessions of remote work	2 sessions of remote work
Experiment Design	Two factors (2 methods, 4 facets))	Two factors (2 methods, 2 facets))
Experiment Group	visual vs textual	visual vs textual
Variables	EFFECT, PEOU, PU, ITU	EFFECT, PEOU, PU, ITU

Facet/Method	**Visual**	**Textual**
Network Security	14	15
DB/Web App. Security	15	14

Table 3: Experimental design

while the remaining said they had no working experiences. With respect to knowledge in privacy technologies and regulations, most of the participants had limited expertise. In contrast, they reported an extensive general knowledge of both security technologies and regulations and standards. Participants also reported good general knowledge in requirements engineering.

1.5 Experimental design

We chose a within-subject design where all participants apply both methods to ensure a sufficient number of observations to produce significant conclusions. In order to avoid learning effects, the participants had to identify threats and security controls for different types of security facets of a Smart Grid application scenario. The security facets included Network Security (Network) and Database/Web Application Security (DB/WebApp). For example, for Network Security facet, participants had to identify network security threats like man-in-the-middle attack or DoS attack and proposed security controls to mitigate them.

The participants were randomly assigned to treatments: half of the participants applied first the visual method to network security facet while the second half applied the methods in the opposite order. Table 3 summarizes how the participants has been assigned to the methods.

1.6 Experimental procedure

The experiment was performed during the Security Engineering course held at University of Trento from September 2013 to January 2014. The experiment was organized in three main phases:

Training. Participants were given a tutorial on the Smart Grid application scenario and a tutorial on visual and textual methods of the duration of two hours each. Then, participants were administered a questionnaire to collect information about their background and their previous knowledge of other methods and they were assigned to facets based on the experimental design.

Application. Once trained on the Smart Grid scenario and the methods, the participants had to repeat the application of the methods on two different facets: Network and DB/WebApp. For each facet, the participants:

- Attended a two hours lecture on the threats and possible security controls specific for the facet but not concretely applied to the scenario.
- Had 2,5 weeks to apply the assigned method to identify threats and security controls specific for the facet.
- Gave a short presentation about the preliminary results of the method application and received feedback.
- Had one week to deliver an intermediate report to get feedback.

At the end of the course in mid January 2014, each participants submitted a final report documenting the application of the methods on the two facets.

Evaluation. In this phase, the experimenters (the authors of this paper) assessed participants final reports while the participants evaluated the method through questionnaires and interviews. After each application phase the participants answered an on-line post-task questionnaire to provide their feedback on method application. In addition, after final report submission each participant was interviewed for half an hour by one of the experimenters to investigate which are the advantages and disadvantages of the methods. Then, at the end of January each participant gave a presentation summarizing their work in front of the experimenters and an expert in security for Smart Grid. The expert evaluated the quality of the threats and security controls delivered by the participants for the Smart Grid application scenario.

The interview guide contained open questions about the overall opinion of the methods, whether the methods help in identification of threats and security controls and about methods' possible advantages and disadvantages. The interview questions were the same for all the interviewees. The post-task questionnaires include the same questions of the one we administered for our previous experiment which was inspired to the Technology Acceptance Model (TAM) [1]. To avoid that the participants answered on "auto-pilot", 15 out of 31 questions were given with the most positive response on the left and the most negative on the right. The interview guide and the post-task questionnaire are reported in [7].

1.7 Changes to the Original Experiment

The experiment reported in this paper differs from the original experiment in that the participants were asked to work individually than in pairs in order to correlate their performance with their perception of the two methods. In addition, we reduced the focus of security risk assessment only to Network security an Database/Web application security to increase the application time provided to the participants. In fact, in the original experiment, participants reported that the time for methods application was short. The main differences are reported in Table 2.

References

[1] F. D. Davis. Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, pages 319–340, 1989.

[2] EUROCONTROL. *EATM, ATM Security Risk Assessment Methodology, Edition 1.0*, May 2008.

[3] ISO/IEC. *31000:2009 – Risk Management.* 2009.

[4] K. Labunets, F. Massacci, F. Paci, and L. M. Tran. An experimental comparison of two risk-based security methods. In *Proc. of ESEM '13*, pages 163–172, 2013.

[5] M. S. Lund, B. Solhaug, and K. Stolen. A guided tour of the coras method. In *Model-Driven Risk Analysis*, pages 23–43. Springer, 2011.

[6] D. L. Moody. The method evaluation model: a theoretical model for validating information systems design methods. In *Proc. of ECIS '03*, pages 1327–1336, 2003.

[7] UNITN. Experiment website. `http://securitylab.disi.unitn.it/doku.php?id=seceng-course-exp-2013`.

[8] C. Wohlin, P. Runeson, M. Hst, M. C. Ohlsson, B. Regnell, and A. Wessln. *Experimentation in software engineering*. Springer, 2012.