Assessing the effect of learning styles on risk model comprehensibility
A controlled experiment (short paper)

Labunets, Katsiaryna; Condori-Fernandez, Nelly

Publication date
2019

Document Version
Final published version

Published in
CEUR Workshop Proceedings

Assessing the effect of learning styles on risk model comprehensibility: A controlled experiment (short paper)

Katsiaryna Labunets
Delft University of Technology, the Netherlands
k.labunets@tudelft.nl

Nelly Condori-Fernandez
Universidade Da Coruna, Spain
n.condori.fernandez@udc.es
Vrije Universiteit Amsterdam, the Netherlands
n.condori-fernandez@vu.nl

Abstract

This paper presents the design of an experimental study and plan for the conduction of a live study with the participants of the REFSQ2019 conference. The study aims to evaluate the effect of learning styles on risk model comprehensibility throughout a controlled experiment. We combine the baseline experiment designed and conducted by one of the authors to assess the comprehensibility of graphical and tabular security risk models with the questionnaires proposed by Soloman and Felder to measure learning style of people. This study will contribute to the state-of-the-art by looking into the effect of learning styles on the communication of security requirements to the stakeholders and whether an appropriate modelling notation type would help to improve risk model comprehensibility.

1 Introduction

There are different learning styles (LSs) among people that could affect how they are susceptible to visual or natural language representations. Having a good match between the LS of the decision maker and the representation could lead to a better understanding of information communicated with that person. This topic is essential for security and software engineering field as the outcomes of security risk assessment have to be communicated mostly with people without a security background and have to be easy to understand (e.g. decision makers at the strategy level).

Although several empirical studies were conducted in requirements inspection to investigate the LSs of individual inspectors, there is not yet enough evidence regarding the effect of LSs on the comprehensibility of risk modeling notations.

In: A. Editor, B. Coeditor (eds.): Proceedings of the XYZ Workshop, Location, Country, DD-MMM-YYYY, published at http://ceur-ws.org

The motivation to conduct the study is twofold. Firstly, we are eager to replicate an experiment (conducted by the first author), with the purpose of investigating the effect of LSs, factor that had not been considered in the baseline experiment [8]. Moreover, as that experiment involved undergraduate students, the setting of the replication within REFSQ would be ideal for involving participants from industry and senior researchers from the academia. Secondly, we would like to corroborate whether a theory from Cognitive psychology [5] that individuals *"have different strengths and preferences in the ways they take in and process information"* is applicable in the context of security risk analysis. This theory was initially proposed to understand LSs in the context of engineering education. The context of our study is similar to Felder and Silverman's theory [6] as stakeholders have to learn information documented in security risk models. Also, we would be able to get a better understanding on the need of matching person characteristics (e.g. cognitive styles, skills) and job task requirements. According to Sims [12], having such matching should increase personal satisfaction and job performance as well as organizational effectiveness.

2 Study Design

2.1 Goal and Research questions

Based on the Goal Question Metric template by Basili [1], we define the goal of our study as follows:

> Our experiment aims to *analyze* risk model in graphical and tabular representations *for the purpose* of assessing the effect of LSs on model comprehensibility *with respect to* the extraction correct information about security risks *from the viewpoint of* the decision-maker *in the context of* industrial practitioners and researchers attending REFSQ 2019 conference.

From this goal, we derive the following research question:

RQ$_1$ *What is the effect of LSs on the comprehensibility of risk models?*

Correspondingly, we define our alternative experimental hypothesis:

H1a: *The participants using a representation that matches their LS will have a better level of comprehension of information in risk model comparing to the participants that does not have match between LS and risk modeling notation.*

2.2 Type of study

To investigate this research problem, we propose to conduct a controlled experiment.

2.3 Relevance of study for research and/or for practice

The results of the proposed study have potential interest to both industrial practitioners and researchers. First of all, it is relevant to industrial practice as it aims at investigating the applicability of Index Learning Style (ILS) [13] to profile decision-makers and the effect of LS on the understanding different security risk modeling notations. The outcomes could potentially lead to recommendations on how to choose an appropriate representation for better security requirements communication. From an academic perspective, this study could show another critical direction in the assessment of modeling notations. It might be the case that the notation designers must take into account stakeholders' LSs. We are going to find out if this is an essential factor for the design of notations.

2.4 Variables and Metrics

We identified two type of variables:

1. *Response variables:* Level of comprehensibility.
2. *Factors:*
   - Learning style (LS), which will be measured using the ILS [13]. The ILS is an online questionnaire that contains 44 questions across four LS dimensions (Sensing/Intuitive, Visual/Verbal, Active/Reflective, Sequential/Global).

Table 1: Experimental Design

Group	Part 1	Part 2
Group 1	OB + Tabular	HCN + Graphical
Group 2	OB + Graphical	HCN + Tabular
Group 3	HCN + Tabular	OB + Graphical
Group 4	HCN + Graphical	OB + Tabular

- Modeling notation, which will be two types: graphical based on CORAS language, and tabular based on NIST 800-30 standard.
- Online Banking (OB) and Health Care Network (HCN) application scenarios are used to control the possible learning effect between experiment parts.

2.5 Population of interest

The intended subjects of our study are industrial practitioners and researchers who play the role of decision makers. No prior background in security or requirements modeling is needed. Working experience of at least 2 years is required.

The possible **benefits to the participants** are that they will have a chance to learn two different notations for representing security requirements. They will also get an idea what information present in security risk models and information about their own ILS. The latter can be shared with the participants after the completion of the experiment in order not to bias the data collection process.

2.6 Study design

The goal of our study is to investigate if there is a synergy between LSs and representation types and what is its effect on the level of comprehension of risk models. Therefore, we chose a within-subject design where the participants complete the comprehension task using both risk modeling notations for two different application scenarios (OB and HCN). This experimental design will allow us to compare the level of comprehension of both types of modeling notations by participants with different LSs. To control the effect of scenarios and modeling notations, we will randomly assign participants to one of four treatment groups described in Table 1.

2.7 Instrumentation

To collect information about participants demographics and background we will ask questions regarding their age, gender, level of English, education degree, working experience, working domain, if they have any experience in security and privacy. We will also ask them to self-evaluate their level of expertise in the relevant areas like requirements engineering, security and privacy technologies and regulations, graphical modeling languages, risk assessment, and application scenario domains.

To identify the LSs of participants, we will use ILS questionnaire [13]. This questionnaire was also used to study the effect of LS on the inspections of requirements artifacts [7] and software [2].

To measure the level of comprehension of risk models, we will use the comprehension questions developed by one of the authors of this proposal and used in her previous studies on risk model comprehensibility [8, 10]. The comprehension task will have six questions about information represented in the model. The questions between two parts of the experiment will be similar regarding the cognitive task to be done and expected response.

The risk models of OB and HCN scenarios were developed with the help of the authors of CORAS language and based on realistic application scenarios developed in collaboration with industrial partners.

2.8 Experimental Procedure

To participate in the experiment, the participants will need to use their laptops. The experimental procedure is the following:

10 min **Introduction:** An introductory briefing to explain participants the high-level goal of the study, task, and what they can expect during the experiment.

3-5 min **Informed consent:** The participants should read the study informed consent and provide their agreement to participate in the experiment.

Table 2: Statistical Test Selection

Comparison Type	Interval/Ratio (Normality is assumed)	Interval/Ratio (Normality is not assumed), Ordinal
2 paired groups	Paired t-test	Wilcoxon test
2 unpaired groups	Unpaired t-test	Mann–Whitney test
3+ matched groups	Repeated-measures ANOVA	Friedman test
3+ unmatched groups	ANOVA	Kruskal–Wallis test

10-15 min **Pre-task:** The pre-task questionnaire will collect demographic and background information about participants and profile them based on the ILS. After this questionnaire, the participants will be randomly assigned to one of four groups presented in Table 1.

5-7 min **Training Part 1:** The participants will have to watch a short video tutorial about the notation that they were assigned and application scenario.

22 min **Application Part 1:** The participants had to review the appointed risk model and answer six comprehension questions. Participants have 20 minutes to complete the task after which they were automatically advanced to the next page. An image of corresponding risk models will be built in on the top of the task page and protected from downloading or opening in another tab in the browser. The tutorial on notation and scenario are provided at the beginning of the task and can be downloaded. After finishing the task, participants fill in a post-task questionnaire.

5 min **Training Part 2:** The participants will have to watch another video tutorial about the second notation and another application scenario.

22 min **Application Part 2:** The participants have to complete similar task as in part 1 but using another notation and application scenario. After finishing the task, participants fill in a post-task questionnaire.

In total the experiment will take up to 90 minutes.

Evaluation: After getting the results, researchers check the responses and mark correct and wrong answers to each comprehension question based on the predefined list of correct responses.

3 Plan of Data collection and analysis

Based on the metrics and instruments used in the experiment, we plan to collect the following data: i) demographics and background data; ii) participants profiles based on ILS; iii) responses to the comprehensibility questions; and iv) responses to post-task questionnaires. For the research hypotheses testing we will use two-way ANOVA or permutation test for two-way ANOVA in case the assumptions of the ANOVA are violated for our samples. To investigate the effect of particular LSs on the level of comprehensibility we select appropriate test based on Table 2 (a short version of Table 37.1 from [11, Chap. 37]). We will also control the effect of co-founding factors (e.g., participants' background, level of English, etc.) on the results in order to be sure that the observed effect is due to the treatments.

4 Threats to the validity and Ethical issues

This section discusses the new threats that were not discussed in the baseline experiment reported in [9].

4.1 Construct validity

It refers to how well the ILS questionnaire measures the learning style of an individual. As we used an instrument empirically validated [5], this threat is mitigated. The ILS will be automatically calculated and reported once the participant completes the experiment.

4.2 Internal validity

The causal relation between the type of learning style and the different notations used for representing risk models could threat internal validity. We mitigated it by adopting a within-subject design and asking participants to complete comprehensibility task with both types of risk models.

4.3 External validity

It refers to the extent to which the results of a study can be generalized to other settings. In our live study, the heterogeneity of subjects (e.g., participants background and experience) would contribute to the external validity of our research. However, this heterogeneity could also bring greater variability in measures affecting the conclusion validity. To reduce this threat, we consider involving only participants with at least 2 years of working experience.

4.4 Ethical issues

The experiment will be implemented using one of existing survey platform (e.g., Qualtrics) and, at the first page, the participants will have to read information about the experiment and privacy statement and give their consent to participate in the study. The participation in the study will be anonymous and volunteer. Therefore, no harm to the participants is present.

5 Publicity and dissemination plan

To make our study public and attract more potential participants we plan to use the social networks of REFSQ2019 (e.g., Twitter, Facebook) and mailing lists. We will ask organizers to help with spreading information about our study, e.g., by including the flyer about the study in the REFSQ2019 participant's package.

The summary of preliminary results will be communicated with attendees in the form of a short presentation on the last day of the conference. The final results and its discussion will be published as a research paper and submitted to one of the appropriate venues either a conference (e.g., ER, ESEM, MODELS, REFSQ, CAISE) or journal (e.g. Journal of Systems and Software).

6 Proposers's bio

Katsiaryna Labunets is a postdoc at the Technische Universiteit Delft (the Nethrlands). She has a significant experience in designing and organizing controlled experiments. She conducted more than 15 experiments of different duration from 1 hour up to 4 months with up to 60 participants. In her research Katsiaryna uses different techniques for collecting quantitative and qualitative data. For the data analysis she is proficient in statistical hypothesis testing (in R) for quantitative data, and grounded theory analysis for qualitative data. The results of the experiments conducted by Labunets have been published in the conferences like ESEM [8] and REFSQ [9], and in the EMSE journal [10].

Her research focus is in investigation of existing doubt if current security methods work and worth to adopt. In particular, she studies the comprehensibility of tabular and graphical notations for representing risk models.

Nelly Condori-Fernandez is an assistant professor at the Universidade da Coruna(Spain) and research associate of the Vrije Universiteit Amsterdam (the Netherlands). Her main empirically-driven research focuses on topics related to quality requirements prioritization and requirements validation. She has a particular interest in applying Human Computer Interaction technologies to support requirements engineering activities. Her research interests also include software sustainability design and assessment with special emphasis on social and technical aspects. She has executed various type of empirical studies and published in conferences like REFSQ, ESEM, EASE and journals as JSS and IST. Nelly has also conducted studies as part of the Live Study track in different editions of REFSQ (e.g. [3, 4]).

References

[1] Basili, V.R., Caldiera, G., Rombach, H.D.: The goal question metric approach. In: Marciniak, J.J. (ed.) Encyclopedia of Software Engineering, vol. 1. John Wiley & Sons (1994)

[2] Carneiro, G., Laigner, R., Kalinowski, M., Winkler, D., Biffl, S.: Investigating the influence of inspector learning styles on design inspections: Findings of a quasi-experiment (2017)

[3] Condori-Fernandez, N., Daneva, M., Wieringa, R.: A survey on empirical requirements engineering research practices. In: Proceedings of the Workshops RE4SuSy, REEW, CreaRE, RePriCo, IWSPM and the Conference Related Empirical Study, Empirical Fair and Doctoral Symposium. pp. 282–295. ICB-Research Report No. 52 (2012)

[4] Condori-Fernandez, N., Lago, P.: Characterizing the contribution of quality requirements to software sustainability. J. Sys. Soft. **137**, 289 – 305 (2018)

[5] Felder, R.M., Spurlin, J.E.: Applications, Reliability, and Validity of the Index of Learning Styles. Intl. J. of Engineering Education **21**(1), 103–112 (2005)

[6] Felder, R.M., Silverman, L.K., et al.: Learning and teaching styles in engineering education. Engr. Education **78**(7), 674–681 (1988)

[7] Goswami, A., Walia, G., McCourt, M., Padmanabhan, G.: Using eye tracking to investigate reading patterns and learning styles of software requirement inspectors to enhance inspection team outcome. In: Proc. of ESEM 2016. p. 34. ACM (2016)

[8] Labunets, K.: No search allowed: what risk modeling notation to choose? In: Proc. of ESEM 2018. p. 20. ACM (2018)

[9] Labunets, K., Massacci, F., Paci, F.: On the equivalence between graphical and tabular representations for security risk assessment. In: Proc. of REFSQ 2017. pp. 191–208. Springer (2017)

[10] Labunets, K., Massacci, F., Paci, F., Marczak, S., de Oliveira, F.M.: Model comprehension for security risk assessment: an empirical comparison of tabular vs. graphical representations. Empir. Soft. Eng. **22**(6), 3017–3056 (2017)

[11] Motulsky, H.: Intuitive Biostatistics: A Nonmathematical Guide to Statistical Thinking. Oxford University Press, New York, USA (1995)

[12] Sims, R.R.: Kolb's experiential learning theory: A framework for assessing person-job interaction. Acad. Manage. Rev. **8**(3), 501–508 (1983)

[13] Soloman, B.A., Felder, R.M.: Index of learning styles questionnaire. NC State University. Available online at: http://www. engr. ncsu. edu/learningstyles/ilsweb. html (last visited on 14.05. 2010) **70** (2005)