Towards Security Requirements: Iconicity as a Feature of an Informal Modeling Language

Alexandr Vasenev[1], Dan Ionita[1], Tommaso Zoppi[2], Andrea Ceccarelli[2], and Roel Wieringa[1]

[1] Services, Cybersecurity and Safety Group, University of Twente
{a.vasenev,d.ionita,r.j.wieringa}@utwente.nl
[2] Resilient Computing Lab, University of Florence
{tommaso.zoppi,andrea.ceccarelli}@unifi.it

Abstract. Self-adaptive systems need to be designed with respect to threats within their operating conditions. Identifying such threats during the design phase can benefit from the involvement of stakeholders. Using a system model, the stakeholders, who may neither be IT experts nor security experts, can identify threats as a first step towards formulating security requirements. To support it, the modeling language might possess adequate features to support this task. This paper investigates how iconic signs as a feature of an informal modeling language can contribute to eliciting security requirements by non-experts. Taking urban grid as a case, we relate benefits and specifics of using iconic signs to the two modeling challenges: i) reducing the cognitive complexity required to understand and model a system by non-experts, and ii) facilitating the threat identification activity using a system model. Outputs of three experiments suggest that iconic signs do assists in addressing the challenges.

Keywords: Requirements elicitation and analysis, Cyber-physical networks, Security requirements, Electrical network, Smart Grid, Experiments

1 Introduction

Modern cyber-physical systems, such as smart grids, should account for the context in which they operate to ensure the continuous service delivery. In principle, designing complex systems demands that multiple stakeholders are directly involved [1]. This might include identifying threats to the system as a part of the security engineering process. On a larger scale, security engineering includes threat modeling, security requirements, and development of security mechanisms [20].

Identifying possible misuse cases of a system – as a step toward threat modeling and then formulating security requirements – leads to earlier focus on

security. In case of self-adapting systems, a list of identified threats can also be later used to consider later how the system should react to specific threats. For instance, adaptation patterns (see, e.g., [23]) can be devised in connection to a particular threat or a threat group. Constructing a list of threats that a self-aware system should account for is challenging, especially if stakeholders with little modeling and security background are involved.

This paper follows the paradigm that requirements engineering starts with problem identification and needs input from stakeholders. The involvement of stakeholders' encourages their creativity and invites them into discussion even if they lack significant technical expertise [2]. Still, the task of eliciting requirements from stakeholders can be complicated due to their different backgrounds [3].

System models can aid in the communication between stakeholders and system architectures. It can hardly be expected that stakeholders, who are concerned with proper functioning of complex adaptive systems, possess significant expertise in modeling (e.g., DFD) or using security-related approaches (e.g., UMLsec). Importantly, modeling notations "should be palatable to the users" [4]. This challenge is closely related to usability of requirements engineering (RE) approaches in general. Although this topic didn't receive significant attention yet, RE community is increasingly concerned about making approaches "usable not only for requirements engineers, but also to stakeholders, with their diverse backgrounds and needs" [5]. Thus, it is desirable to study (and provide empirically-based suggestions) what aspects can impact understanding and effectiveness of non-specialists involved in modeling security requirements, including its threat identification as the first step.

Even though a number of researchers concentrated on visual notations (see, e.g., a seminal work [6] on the topic), several questions on their role in requirements engineering of complex systems are still open. Specifically, the question is still open whether an icon-based representation can indeed assist threat identification by stakeholders and influence perception of stakeholders about this task. Moreover, to our best knowledge, little empirical studies are published on this topic. This paper makes initial steps towards answering this questions. While it doesn't claim statistically significant results, it provides initial support for the argument that using iconic informal languages can assist in eliciting security requirements. For this, we draw on advances in conceptual modeling [8], [9], [10], a well-developed topic in the information systems domain, which often deals with large-scale systems.

We study how iconicity [7] can contribute to modeling and threat identification by non-experts. *Iconicity* is seen as a relation of resemblance or similarity between the two main aspects of a sign: its form and its meaning. Herewith, we consider iconicity to be related to two modeling challenges (MC):

- MC1. To support the reduction of the cognitive complexity required to understand and model a system;
- MC2. To facilitate the threat identification activity using a system model.

For the purpose of this research, we take an urban electricity grid as an example of an adaptive cyber-physical system. The grid model represents city-level grid components (e.g., a power substation, hospital) and connections between them. Such a model is similar to a UML deployment diagram, and consists of: i) nodes as modelling elements that represent the system components and ii) links among the nodes.

After reviewing relevant background in the next section, the paper introduces the methodology used, presents results of three experiments, discusses them, and concludes with future work.

2 Background and Motivation

Often, different stakeholders should collaborate to ensure that a system will deliver desired services. This is particularly relevant for *Critical Infrastructures (CIs)*. Typically, CIs operate in complex social, economic, and technical contexts that imply collaborations between a number of stakeholders. The focus of risk management in CI concerns threats to safety and security. For comparison, risk management in the insurance, engineering, and finance domains aims at protecting against financial losses. All these aspects justify considering an urban electricity grid, which is a specific CI, as a particular relevant case for studying the security requirements elicitation process.

Being complex systems, urban electricity grids need to re-act to changes in their environment. In normal operation mode a number of elements are highly interconnected. At the same time, grid should be ready to adapt to for rapid changes in real-time. For instance, a part of the grid (including both power and ICT systems) might be able to became an autonomously operating island (or a microgrid) to prevent cascading failures. Proper identifying and modeling of threats is critical to devise adequate security mechanisms.

Numerous grid stakeholders possess specific (tacit) knowledge relevant to ensuring proper functioning of the grid. *City managers* might have significant expertise in daily administrative operations. They might ensure how renewables are related to renewable energy-related landscape features and to the reduction of greenhouse gas emission [11]. *Grid operators* are responsible for day-to-day functioning of the infrastructure and should consider how the interplay between the urban form and solar energy inputs should be taken into account [12]. *Specialized agencies* might have expertise in contingency planning and risk assessment. Still, each individual stakeholder might lack an overall picture of city development strategies, contexts, and specifics of threat landscape. This calls for the need for them to work together.

Several solutions exist to assist stakeholders in identifying risks and threats to the grid, as shown in the next subsection. However, it is unclear how representations of a modeling language can assist in modeling a system (*MC1*) and identifying threats to it (*MC2*).

2.1 Risk Assessment Approaches

The need for adequate risk management when considering CI is highlighted by government agencies [13]. Specifically, identifying threats to the grid is an important step towards supporting definition of adequate responses [14]. This step should account for the involvement of non-experts.

Often, CI-focused risk assessment methodologies are built around three core tasks that start with threat identification: i) identification and classification of threats, ii) identification of vulnerabilities and iii) evaluation of impact. For instance, Risk Assessment (RA) for CIs, according to *U.S. Department for Homeland Security*, should always start with obtaining *a clear and agreed view of the infrastructure* from all public and private partners. This shared view is used to assess the relevant risks, in terms of *threats*, *vulnerabilities*, and *impacts*[13]. Thus, supporting solutions to construct system models and identify threats is of significant importance.

Several tools were developed to support critical infrastructure analysis. Primarily, they address the policy maker point of view. An example is the CARVER tool (Criticality Accessibility Recoverability Vulnerability Espyability Redundancy) [15]. Similarly, the tool called Critical Infrastructure Modelling Simulation (CIMS) aims at model-based CI disruption simulation. This provides policy makers with decision-support when faced with threats and natural disasters [16]. One more example is a tool [17] oriented to involve non-technical users. Altogether, the focus of these tools rely emphasizes the role of involving non-technical experts to consider threats to the system.

2.2 Iconic and Non-Iconic Signs

It is desirable that (informal) modeling languages can contribute to reducing the cognitive complexity of a modeling task (*MC1*) and identifying threats to a system (*MC2*). Together, the two challenges concern investigating utility and usability of different representation of a grid modeling language for the threat identification task.

Cognitive theories support the argument that a less suitable format of the model representation (the signs used) can hamper understanding of non-experts. These theories particularity stress the importance of having an intuitive and understandable representation of concepts. *Cognitive fit* theory, for instance, suggests that the cognitive load is reduced when a representation matches the problem [18]. Earlier, it was found that visually recognizable representations speed up creation and improve understandability of multi-layered models, especially when domain experts (who are not modelling specialists) are involved [10]. Yet, it was not studied whether representation of a modeling language can contribute to the threat identification task. This task is linked to the modeling exercise, as well concerns security requirements at large.

This paper describes a study on the basic differentiation between using iconic and symbolic signs. *Iconic* signs visually resemble the concepts that they represent, whereas *symbolic* signs are arbitrary and the relationship with the concept

they represent is purely conventional. We investigate the effect of utilizing either symbolic or iconic signs for the modelling and threat identification tasks given the same list of modelling constructs. Noticeably, we go beyond a simple consideration that iconic signs are beneficial for understanding concepts, as we study the role of signs within a modelling language to identify threats to a system. The experiment design, devised for this paper, accounts for these challenges, as shown next.

3 Methodology

To study how iconicity influences modeling and threat identification, we designed and conducted three experiments. These experiments are interrelated as shown in Fig. 1. The combination of three experiments covered different combinations of possible usage of the iconic modeling language. The focus of the second and the third experiments concerned modeling a system and identifying threats to a system accordingly. Together, the system of experiments dealt with both modeling and threat identification steps. With respect to evaluation criteria, we studied perception of users and compared amounts of threats identified by different groups.

The participants were provided with either iconic or textual grid elements for experiments 1 and 2. Constructed grid models (either iconic or textual ones) together with a generic threat list formed the input to experiment 3. The configuration of experiments formed a structure that assessed the utility of iconic models, influence of iconicity to model the grid, and its role in identifying threats. As real experts were unavailable, we used a sample of BSc, MSc and PhD students. Experiments 1 and 2 were conducted at University of Twente (UT) during the CuriousU summer school. Later, Experiment 3 took place at University of Florence (UNIFI). This section describes the set-up in detail.

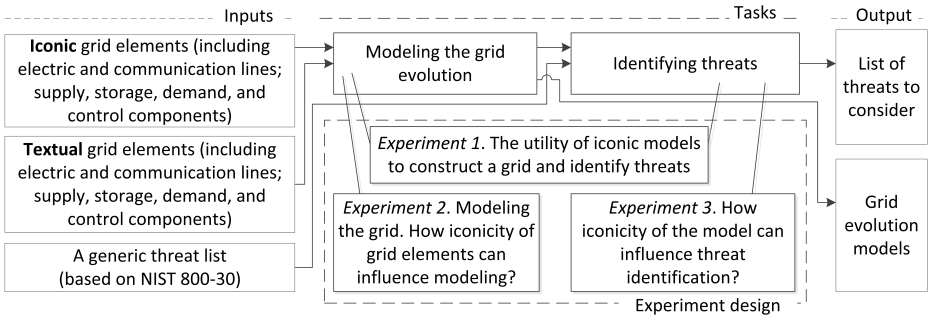

Fig. 1. Outline and relations between the three experiments.

The experiments tackled challenges *MC1* and *MC2* as shown in Table 1. Also, the table describes sample populations, modeling tartgets, and treatments of the three experiments.

Table 1. Characteristics of the experiments

	Experiment		
	1	**2**	**3**
Challenges tackled	MC1, MC2	MC1	MC2
Sample Population	2 groups of 6-8 participants		2 groups of 3 participants
Modeling Target	Infrastructure of the grid on UT campus		Model of UNIFI area
Treatment	Design a prototype using provided software tools		List of threat occurrences in the given scenario

The intended target of generalization of the three experiments are city-level stakeholders with no specific experience on modeling and threat identification. While students are not representative city planners (and we acknowledge that it somewhat weakens evaluation efforts), the outcomes of the experiments are produced by general cognitive mechanisms (such as cognitive fit) which are shared by both groups. Furthermore, both students and city planners are unlikely to possess knowledge or experience with regard to critical infrastructure modelling tools or threat identification techniques. We did not control for pre-existing knowledge of the students, but to enhance external validity, participants were firstly introduced to typical infrastructure components of grids simulating the basic knowledge that city planners might have.

Participants, supplied with specific materials, joined one of the three experiments. Each experiment consisted of a task performed by two groups. Members of each group worked collectively on their task. After the completion of task, the participants filled-out their questionnaires individually. While one can debate whether the individual questionnaire responses are fully independent, each task implied the idea of collaborative work. In this connection, we were interested if a specific representation (possibly, by supporting interactions) influenced perceptions of individuals after they completed a collaborative task.

Due to self-forming, the group sizes were not equal. Still, the amount of people in each of the groups (within a single experiment) did not deviate largely (see Table 1). The difference can potentially be seen as a threat. However, as the experiments involved participants of a summer school and visiting students, they (to our best knowledge) did not have significant previous experience of working with each other. Thus, we did not limit the way how the groups were formed. While we didn't collect group profiles, participants within each experiment shared a similar level of education. In each of the three experiments, the treatments were randomly allocated to the groups involved.

Experiment 1 focused on whether modeling a grid (*MC1*) and identifying threats (*MC2*) can be performed within a comparable time interval by using iconic or symbolic modeling constructs. For this, the groups were tasked to construct a model how they imagine the campus grid of the *UT* in 5–10 years. Potential threats to the validity of this experiment that we could not control are pre-existing security or safety knowledge and experimenter expectancy (as the exercise was supervised). However, we attempted to ensure treatment and measurement validity by running the two sessions in parallel provided each group with the same tools (MS Visio) and instructions (handouts). Two supervisors involved in the experiment were allowed only to answer questions strictly related to the threat lists.

Experiment 2 concerned only with the modeling task and did not cover the threat identification step. It investigated whether iconicity of the modeling language would influence modeling changes in the system and understandability (*MC1*). After performing the task, the participants filled in a questionnaire formed by 4 questions to document their perception of difficulty and success of the modeling task. Questionnaire design asked for a score from 1 to 5 to each question following a psychometric semantic differential scale to reduce acquiescence bias [19]. The questions were as follows:

- E2Q1. "How would you describe the difficulty of the task you just completed? *Rate from 1 (Very easy) to 5 (Very Difficult)*;
- E2Q2. "How satisfied are you with the tools provided to complete the task? *Rate from 1 (Not Satisfied) to 5 (Very Satisfied)*;
- E2Q3. "How would you rate the amount of time it took to complete the task? *Rate from 1 (Very little time) to 5 (Too long)*;
- E2Q4. "How much do you agree with the final version of the model? *Rate from 1 (Don't agree) to 5 (Fully agree).*

This second experiment was conducted under stricter conditions: supervisors were not allowed to assist the modelers. Participants answered printed questionnaires immediately after the task. However, group dynamics could have influenced the measurement validity. For instance, one can assume that some participants might have reported lower agreement or perceived the task as more difficult due to intra-group personality or skill mismatches. The experiments did not investigated either of these aspects. Nevertheless, as the groups were formed from a pool of participants with similar education experiences, we expect that influences of these aspects were limited. Another threat to validity to the second (as well as to the first) experiment is that both groups worked in a single, although very large room. To counter it two supervisors tried to limit cross-group interaction.

Experiment 3 explicitly dealt with identifying threats to a grid. It concentrated on how participants relate an iconic or symbolic grid model to a generic threat list. It was designed to understand how the iconicity feature of a model influences the ability of non-experts to perform an effective – complete, precise, and accurate – threat identification task (*MC2*). After defining two groups of 3

students at *UNIFI*, we asked the participants to identify all the possible threat occurrences of a given modeled scenario considering a reference threat list [21]. We supplied all participants with the same scenario, described either in iconic or symbolic signs. The *independent variable* (iconicity of constructs), thus, was thus similar as in Experiment 1 and Experiment 2. See Table 1 for details.

The obtained threat lists were compared with a list provided by an expert from UNIFI to assess the completeness of students' lists. Also, the participants filled in the following questionnaire:

- E3Q1. "How would you describe the difficulty of building the list of threats? *Rate from 1 (Very easy) to 5 (Very Difficult)*;
- E3Q2. "Was the graphical/symbolic description enough to complete the task? *Rate from: 1 (Unnecessary) to 5 (Very Useful)*;
- E3Q3. "Did you feel that additional software supports were needed? *Rate from: 1 (No) to 5 (Yes, I was lost)*;
- E3Q4. "How would you rate the amount of time it took to complete the task? *Rate from: 1 (Very little time) to 5 (Too long)*;
- E3Q5. "Do you feel that the list you provided is complete? *Rate from 1 (Very poor list) to 5 (Very complete list).*

Afterwards, we asked the students to anticipate how the threat identification exercise would be if they would have the model described in another way (iconic for *Group 2* and symbolic for *Group 1*). By doing so, we aimed to collect perceived benefits of using alternative description of a scenario. The threat identification exercise was not repeated. To differentiate between the two questionnaires filled out, the rest of the paper refers to the "perceived benefits" questionnaire as *Experiment 3b*, while the initial survey is referred to as *Experiment 3a*.

4 Experiments and Findings

Input to the first and the second experiments included lists of i) generic threats to grid components and ii) either an iconic or a symbolic list of grid components to build an urban grid. The latter input was organized as a template in a MS Visio file. In the third experiment, the students were supplied with a list of generic threats and with either an iconic or symbolic model. The provided model was similar in complexity to those obtained during the first two exercises.

Iconic modeling constructs are described in [21] and form pairs (icon-name). Some icons are included in Fig. 2. In the *symbolic* template, the modeling constructs were presented only by their names (e.g., 'power substation', 'wind farm', and 'hospital'), without icons.

4.1 Experiment 1

This experiment aimed to consider the utility of the provided language to model the grid and identify threats to it. The main task was to create grid models

(see, e.g., Fig. 2). Also, we asked participants to identify threats relevant to particular steps of the grid development (using a generic list of possible threats, as described in [21]) and relate evolution to threat sources (in terms of their capability, intent, and targeting characteristics. This secondary task investigated whether participants can meaningfully relate the grid structure they constructed with the idea of threat modeling. By doing so, we intended to position the task of threat identification in the context of security engineering. Altogether, we aimed at investigating whether constructing a grid model and identifying threats to it can be feasible for both iconic and symbolic groups.

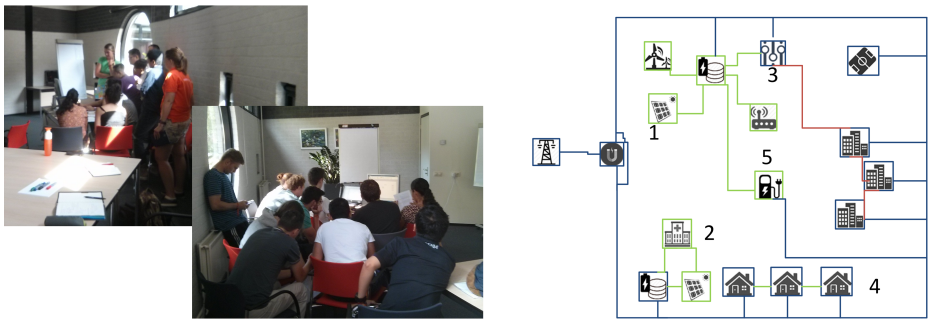

Fig. 2. Experiment 1 running and the grid structure constructed by one of the groups (numbers in the figure indicate steps when new components are introduced).

Main Findings. An interesting finding of this experiment was that the iconic group decided to proceed with modeling the grid in MS Visio directly, while another group started to draft their plans on a whiteboard and paper sheets. We did not anticipate that groups would utilize alternative media when confronted with non-iconic notations. An explanation could be that in this case a lack of iconicity eliminated perceived benefits of using a software-modelling tool, while the flexibility afforded by free-hand drawing led to the use of whiteboard. This potentially points out that the notation of a modelling language can directly impact the modelling process. Both groups were capable to construct grid models and identify a comparable number of relevant threats, despite their previous lack of experience with this task. It suggests that the both representations, as well as the language, can be used for relating components to threat sources.

4.2 Experiment 2

This experiment concentrated on obtaining initial quantitative data whether modeling using software tools with iconic signs is perceived by non-experts as more understandable compared to modeling with non-iconic signs. Similar to Experiment 1, two groups of ten students each were asked to construct models of a smart future university campus. Afterwards, we collected four questionnaires

from the group that used iconic signs (*Group 1*) and seven questionnaires from the other group (*Group 2*).

Main Findings. Table 2 describes the collected data. The members of *Group 2* found the task more difficult (by 64%) and were less satisfied with the tool to model the infrastructure (24%). The E2Q1 answers from the two groups differ significantly and their confidence intervals do not overlap. It highlights difficulties that the students from *Group 2* encountered during modeling the future grid. The replies to E2Q3 and E2Q4 are less illustrative: while being comparable, they deviate largely.

Table 2. Experiment 2: Averages and standard deviations

	Iconic signs (Group 1)	**Symbolic (Group 2)**
E2Q1	Avg 2,0 (Std 0,7)	3,3 (0,5)
E2Q2	3,8 (0,4)	2,8 (0,8)
E2Q3	2,5 (1,1)	2,9 (0,5)
E2Q4	3,0 (0,7)	3,0 (1,0)

The outcome of this experiment suggests that software-based modelling with iconic signs is perceived as less difficult than when using symbolic signs.

4.3 Experiment 3

The last experiment focused on investigating how an iconic/non-iconic model influences the outcomes of the threat identification task. Two groups each of 3 students participated in the experiment: *Group 1* worked with an iconic description of the grid of the scientific complex of *UNIFI*, while *Group 2* worked with a non-iconic (symbolic) version. Provided with a list of generic threats (a subset of threats 7, 10, 17, 18, 19, 21, 24, 29, 31, 37 of the threat list in Appendix B of [21]), all students built a threat list to the system model. In Table 4 'A' and 'B' letters in the questions distinguish between questionnaires for *Experiment 3a* and *3b*.

Main Findings. Table 3 shows that the amount of valid identified threats is significantly higher for participants who were supplied with the iconic model. *Group 1* members identified 17, 10, and 19 threats. Members from *Group 2* identified 8, 8, and 9 valid threats.

The expert evaluated most of the threats identified by the students as being valid. Some threats, e.g., "conduct physical attacks on organizational facilities", were commonly identified. Some others threats were identified less often ((for instance, only two out of six students identified "conduct attacks using unauthorised ports, protocols and services"). An explanation can be that some threats are difficult to understand (and identify), because they require specific technical knowledge.

The *Iconic* group reported less difficulty (E3AQ1) and more satisfaction of the results (E3AQ5). Also, they were indicated (E3AQ3) that additional software

Table 3. Experiment 3: Averages and standard deviations

Questions	Iconic (Group 1)	Symbolic (Group 2)
Experiment 3a (answers 1 to 5)		
E3AQ1	Avg 3.0 (Std 0)	3.7 (0.6)
E3AQ2	5.0 (0.0)	4.0 (1.0)
E3AQ3	1.3 (0.6)	2.7 (0.6)
E3AQ4	2.7 (0.6)	3.3 (0.6)
E3AQ5	3.0 (1.0)	2.0 (0.0)
Experiment 3b		
E3BQ1	4.0 (0.0)	3.0 (0.0)
E3BQ2	4.0 (1.0)	4.0 (1.0)
E3BQ3	3.0 (1.0)	3.0 (1.0)
E3BQ4	3.3 (0.6)	2.7 (0.6)
E3BQ5	3.0 (1.0)	2.3 (0.6)
Threats (Amount)		
Identified Threats	15.3 (4.7)	8.3 (0.6)

support is needed less, if compared to the symbolic group. Interestingly, the participants didn't anticipate that employing another representation format can result in a more complete list of threats. E3AQ5 and E3BQ5 answers of *Group 1* both score 3.0. More specifically, there is only a relatively small increase (0.3) in the difference between E3BQ5 and E3AQ5 for *Group 2*.

In summary, all subjects in possession of the iconic model constructed more complete lists of plausible threats compared to their counterparts. It suggests that the threat identification task can benefit from employing an iconic model of a system.

5 Discussion

5.1 Modeling Challenges

MC1: Reduction of cognitive complexity. While *Experiment 1* showed that both notations can be potentially used to identify threats to a system, E2Q1 from *Experiment 2* and to a smaller extent E2Q3 showed that the perceived difficulty of the modeling task slightly decrease when iconic signs are used. Notably, the *Iconic* group was less satisfied with the tools provided (E2Q2). Nevertheless, based on the outcome of the experiments we can argue that the use of iconic signs instead of symbolic ones lowered the cognitive complexity of the task.

MC2: Facilitating threat identification. In general, non-expert users can identify threats to a system regardless of the model's representation (*Experiment 1*). However, if supplied with a readily made iconic models — in contrast to a symbolic one – they performed better (*Experiment 3*) and considered that such the iconic description was completely enough to perform the task.

5.2 Practical Implications

As noted in [20], enumerating threats helps system architects to develop realistic and meaningful security requirements. Thus, this paper contributes to the process of working on security requirements at large.

Specifically, this research provides initial empirical support for claims related to devising and employing means for eliciting security requirements. Our findings hint at high-level suggestions how to approach eliciting security requirements from stakeholders who are less experienced in modeling. In particular: i) using icons for modeling compared to pure text representation of modeling constructs facilitates comprehension of non-experts; ii) iconic models can assist in identifying potential threats by non-experts. We envision that an informal iconic model of a system, such as the one shown in Fig. 2, can facilitate collaboration between stakeholders.

5.3 Limitations

Notes on experiments. Some aspects related to the configuration of experiments should be noted. First, it can be possible that outcomes of the experiments were obtained by pure chance. However, it is the consistency of outcomes of several experiments that points out that using icon-based informal modeling language can be useful to identify threats to a complex system. Second, the experiments were conducted with a low number of participants. Nevertheless, the configuration of experiments was not intended to make statistic-based claims. The potential for generalizations is related to ideas within cognitive theories. More experiments with larger amounts of groups will clearly be beneficial. Third, the experiments were focused on assets-threats connections. We did not account for compliance obligations, raw requirements, security requirements, as well as security measures at large. All these aspects (see, e.g., [22] for a structure of interrelations) are important for security requirements engineering. Investigating the effect of iconicity in connection to other security requirement engineering processes might be a direction for future research. Fourth, the impact of iconicity may be different if the users only identify threats or model and identify threats as two consequent steps. This aspect, as well as the question how qualitative results can be related to quantitative ones in case of threat identification, deserves further studies. Firth, we didn't aim to cover specific expertise of stakeholders. Still, we can expect that stakeholders involved in modeling could have participated in a BSc, MSc, or PhD program. While stakeholders' educational background may be different from the students, critical thinking, analytic, and other skills obtained through their education might be similar. Finally, it is possible that real-world applications might require several modeling sessions, where users over time will become more familiar with construct and their representations. Still, if iterations are rare, stakeholders might need to (re-)familiarize themselves with those elements, similarly to the first time exercise. The outcomes of this research can still be useful in such cases.

Model Quality. In this study the semantics (i.e., correctness and completeness) of the models was not investigated in detail. Also, although an RA expert examined the threats identified by students within *Experiment 3*, we cannot make any claims with regard to the effects of iconicity on the absolute quality of the results. Besides, the "quality" of the identified threats was not part of the evaluation. It is the next steps of the security development process that should account for such a merit. Besides, we did not study how iconicity can explicate tacit knowledge (as experts are needed for this task) and creativity (students were provided with a list of possible threats). Still, we can anticipate that iconic models, due the reduction of cognitive load, can also contribute to these aspects.

Adherence to syntax. We observed that groups with symbolic signs started to freely draw schemes on the whiteboard, thereby reducing possibilities to enforce syntax of the modelling language. In connection to the cognitive fit theory [18], we can explain that another way used to represent information suited the task (and the audiences) better. However, benefits and limitations of using a specific media were not investigated. Possibly, *dual encoding* (illustrating the text corresponding to the components next to their graphical representation) can support efficiently employing different media for modeling.

Choice of signs. Symbolic signs were kept as simple as possible, by using only boxes, arrows and colors. However, the complexity and suitability of iconic signs were not evaluated. It is possible that these icons can be simplified, employ more discriminable symbols, and possess more semantic transparency. Also, this research didn't concern the modeling constructs themselves, as well as portability of the modeling approach to a large-scale scenario. Huge CPS with lots of detail might call for finding a particular level of abstraction. We acknowledge that the set-up of this research accounts for only a fragment of the real world's complexity. It does not investigate how having a very large number of iconic representations can negatively impact human comprehension because of, for instance, similarity across potentially similar elements. We can expect that in such cases modeling languages might benefit from grouping elements. Also, different notations aspects [6] can be applied. This question, next to how the cost of icon design can influence modeling process, was not considered in this paper.

6 Conclusions and Future Work

Eliciting security requirements, as a collaborative process that starts with identifying threats (in other words, misuse cases), should account for inputs from diverse stakeholders. In this paper, we empirically investigated iconicity — a feature of an informal modeling language. We concentrated on identifying threats to an urban grid as a case of cyber-physical system.

Our findings indicate that individuals with little modelling experience do benefit from employing an iconic representations of an informal language. Participants of three experiments perceived iconic models as easier to construct and more understandable. Moreover, participants equipped with the iconic model were capable to point out more threats relevant to the system.

From a requirements standpoint, the findings suggest that iconic representations of informal modeling languages constructs can benefit the threat identification task. This task is an important step of security requirements. Those concerned with developing and employing languages and tools for security requirements can consider employing such notations. Ultimately, the findings can also assist specialists involved in communicating risk assessment and risk management processes to stakeholders. Future work should attempt to replicate the experiments at a large scale, ideally with practitioners.

Acknowledgments This work has been partially supported by the Joint Program Initiative (JPI) Urban Europe via the IRENE project and has received funding from the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement no 318003 (TREsPASS). We thank the students who participated in the experiments. This publication reflects only the author's views and the Union is not liable for any use that may be made of the information contained herein.

References

1. Murer, S., Bonati B.: Managed evolution: a strategy for very large information systems. Springer Science & Business Media (2010)
2. Sindre, G., Opdahl, A. L.: Eliciting security requirements with misuse cases. J. Requirements Engineering, 10(1), 34–44 (2005)
3. Houmb, S. H., Islam, S., Knauss, E., Jrjens, J., Schneider, K.: Eliciting security requirements and tracing them to design: an integration of Common Criteria, heuristics, and UMLsec. J. Requirements Engineering, 15(1), 63–93 (2010)
4. Hickey, A.M.,Davis, A.M..: Elicitation technique selection: how do experts do it? In: 11th IEEE International Requirements Engineering conference (2003)
5. Bombonatti, D., Gralha, C., Moreira, A., Araujo, J., Goulao, M.: Usability of requirements techniques: a systematic literature review. In: 31st Annual ACM Symposium on Applied Computing, Pisa, Italy (2016)
6. Moody, D.: The "Physics" of Notations: Toward a Scientific Basis for Constructing Visual Notations in Software Engineering. J. IEEE Trans. Softw. Eng. 35, 6, 756–779(2009)
7. Aissen, J.: Differential object marking: Iconicity vs. economy. J. Natural Language & Linguistic Theory 21.3 435–483 (2003)
8. Barjis, J.: Collaborative, participative and interactive enterprise modeling. In: Enterprise information systems. Springer Berlin Heidelberg, 651–662 (2009)
9. Hernantes, J., et al.: Collaborative modeling of awareness in Critical Infrastructure Protection. In: 44th IEEE Hawaii International Conference on System Sciences (HICSS) (2011)
10. Ionita, D., Wieringa, R., Bullee, J.-W., Vasenev, A.: Tangible modeling to elicit domain knowledge: an experiment and focus group. In: Conceptual Modeling. Springer International Publishing, 558–565 (2015)
11. Zubelzu, S., lvarez, R., Hernndez, A.: Methodology to calculate the carbon footprint of household land use in the urban planning stage. J. Land Use Policy 48, 223–235 (2015)

12. Amado, M., Poggi, F.: Solar Urban Planning: a parametric approach. J. Energy Procedia 48, 1539–1548 (2014)
13. U.S. Department of Homeland Security.: NIPP Supplemental Tool: Executing a Critical Infrastructure Risk Management Approach (2013)
14. Federal Emergency Management Agency.: A Whole Community Approach to Emergency Management: Principles, Themes, and Pathways for Action (2011)
15. Bennett, B.T.: Understanding, assessing, and responding to terrorism: Protecting critical infrastructure and personnel. John Wiley & Sons (2007)
16. Pederson, P., Dudenhoeffer, D., Hartley, S., Permann, M.: Critical Infrastructure. Idaho National Laboratory (2006)
17. Giannopoulos, G., Filippini, R., Schimmer M.: Risk assessment methodologies for critical infrastructure protection, part I: A state of the art. In: JRC Technical Notes (2012)
18. Vessey, I., Galletta, D.: Cognitive fit: an empirical study of information acquisition. J. Inf. Syst. Res. 2(1) 63-84 (1991)
19. Oddgeir, F., Martinussen, M., Rosenvinge, J.H.: Likert-based vs. semantic differential-based scorings of positive psychological constructs: A psychometric comparison of two versions of a scale measuring resilience. J. Personality and Individual Differences 40.5 873–884 (2006)
20. Myagmar, S., Adam J.L., William, Y.: Threat Modeling as a Basis for Security Requirements. In: Symposium on Requirements Engineering for Information Security (SREIS) (2005)
21. Improving the robustness of urban electricity networks (IRENE) project: Deliverable D2.1 Threats identification and ranking, `http://ireneproject.eu/wp-content/uploads/2016/01/IRENE-D2.1.pdf`
22. Schmitt, C., Liggesmeyer, P.: A Model for Structuring and Reusing Security Requirements Sources and Security Requirements. In: REFSQ Workshops (2015)
23. Menasce, D.A., Gomaa, H., Malek, S., Sousa, J.P: SASSY: A framework for self-architecting service-oriented systems, IEEE Software, 28(6) 78–85 (2011)