Ahab's Leg Dilemma: on the Design of a Controlled Experiment

Luca Sabatucci, Mariano Ceccato, Alessandro Marchetto, Angelo Susi
Fondazione Bruno Kessler
Trento, Italy
{*sabatucci,ceccato,marchetto,susi*}*@fbk.eu*

***Abstract*—To meet stakeholder non-technical background, requirements are often presented by analysts in terms of scenarios. While translating requirements into scenarios, details and over-specifications (called Ahab's Legs) need to be added to make requirements concrete and understandable to stakeholders. Despite the expected benefits that they should convey, Ahab's Legs could disturb the requirement validation session. They can, in fact, distract the attention of stakeholders. Valuable discussion time may be wasted when focusing on irrelevant details rather than on the actually relevant ones.**

In the present paper, we address the Ahab's Leg dilemma and its potential impact on requirement validation sessions. We discuss how to measure the distraction due to Ahab's Legs and what are the possible approaches an analyst can adopt to limit it. Moreover, we present the design of a controlled experiment devoted to measure the impact of Ahab's Legs on requirement validation sessions. In particular, the experiment is meant to (1) estimate the magnitude of the distracting effect, and to (2) assess one of the most promising way to alleviate their negative effect, i.e. by making stakeholder aware of the Ahab's Legs before the validation session.

I. Introduction

The *Ahab's leg dilemma* is a semiotics phenomenon that consists of the need to add details to a story when changing the target media or the communication style, in order to keep the story engaging. This phenomenon has already been discussed by an Italian philosopher, Umberto Eco [1], who observed an example in the famous novel Moby Dick where the main character, Captain Ahab, has a peg-leg. Eco observes that the author, Herman Melville, does not mention whether the peg-leg is the left leg or the right one. However, in the filmic transpositions of Melville book, because of the constraint of the visual media, directors were forced to take a decision about which leg is the whale bone peg. Although the peg-leg is a fundamental part of the story (you cannot imagine any adaptation of the book for which Ahab does not have a peg-leg), knowing which one has no bearing on it. Yet, when the peg-leg is instantiated, this decision may generate a lot of consequences, some of them might be harmless and some might not. As a matter of the fact, different film transpositions took opposite decisions of which leg is the peg one.

Coming back to requirement engineering, we faced a similar problem when we formulated scenarios starting from requirements [5] in a project devoted to implement an automatically supervised medical environment, the ACube[1] project. In our experience, the phenomenon was particularly relevant during the requirement validation phase [7], because the analyst adopted narration as a way to communicate and validate semi-formal requirements with stakeholders. Narrative scenarios were derived from requirements, and the analyst was forced to add details during the translations, in order to instantiate generic requirements into a concrete spatial-temporal context. Although narrative scenarios are an expressive way to represent and communicate requirements to non-technical people (in our case nurses and doctors), we observed that stakeholders sometimes focused their attention on those non-central aspects of the story (Ahab's Legs) that are just due to the translation. This caused waste of valuable time during the session, since comments raised on Ahab's Legs did not impact any part of the actual requirements. In other words an Ahab's Leg potentially represents a source of *distraction* for a stakeholder, who is supposed to provide feedback to the requirements.

The major role of Ahab's Legs, observed during the requirement validation sessions in the ACube project, motivates us in further thinking of this phenomenon. In particular, we would like to investigate the following research questions:

- **Q1**: *What is the actual impact of Ahab's Legs on distracting stakeholders during requirements validation sessions?*
- **Q2**: *Is there a way to to reduce the impact of Ahab's Legs on the stakeholder distraction?*

In other words, we conjecture that the observed negative impact of Ahab's Legs is not limited to the particular ACube context, but it is more general, as it is a matter of course when translating requirements to scenarios for validation purpose. Hence, understanding how to manage the presence of Ahab's Legs becomes a fundamental aspect to consider in order to be able to conduct validation sessions successfully.

In order to verify the conjecture raised during a real project, one opportunity is to conduct repeatable in-lab controlled experiments, as formalized by Wohlin et al. [9]. This would consist in having (artificial) requirement validation

[1]ACube is a large research project funded by the local government of the Autonomous Province of Trento in Italy with the aim of designing a highly technological smart environment to be deployed in nursing homes as a support to medical and assistance staff

sessions where we can control and measure all the relevant variables. In particular, we could change only one dimension on the experiment (i.e. the treatment), while keeping the other controlled, and measure the reaction of stakeholders, so as to study the impact of the Ahab's Leg dilemma.

In this paper we discuss the complex problems connected with the presence of Ahab's Legs in the scenarios used to validate requirements with stakeholders. We would like to understand their impact on the quality of validation sessions and how they should be managed by an analyst in order to limit the most prominent negative effects. Among the most promising approaches to limit the negative impact of Ahab's Legs, we identify one for the empirical validation. We define the design of a controlled experiment devoted to measure the distraction caused by Ahab's Legs during requirement validation sessions and if distraction can be effectively limited when a proper approach is adopted.

The rest of the paper is organized as follows. Section II describes and discusses about the Ahab's Leg dilemma. Section III presents the design of the experiment to investigate about it. Section IV gives an overview of the related literature in the context of requirements validation. Finally, Section V concludes the paper and introduces our future work on this topic.

II. The Ahab's Leg dilemma

In this section we describe the Ahab's Legs dilemma from a practical point of view, starting from an actual example observed during the ACube project.

A. Motivating example

Figure 1 shows the pictorial representation used during a requirement validation session with stakeholders of the ACube project. In the scenario Maria, a patient, is falling in the staircase. The system detects this event and notifies a caregiver by sending a high priority message to her PDA. This scenario is meant to validate a feature of the system, i.e. automatic detection of a dangerous situations and immediate notification to a caregiver, but the kind of sensors (the camera) and the device for notification (the PDA) are still not decided.

The validation session was proceeding in a satisfactory way, until a nurse commented negatively about the adoption of a PDA, shown by the analyst just as an example of device to interact with the ACube system. She complained that a PDA is intrusive, because it would have required to change her working practices, as she had to carry the PDA in her pocket, with battery always load. A very lively discussion began about the possibility of use less invasive devices. However, the purpose of the facilitator was to focus the attention of the group to a different feature to be validated, i.e. *"the system communicates with caregivers with low and high priority signals"*, as the kind of device to use was still not decided, so it was pointless to discuss about it at that stage of the project.

B. The dilemma

Inspired by the Italian philosopher, Umberto Eco, the Ahab's Leg dilemma is defined as the necessity to add more details to the original storyline, because of the different characteristics of the target media (e.g.: textual Vs visual) or because of the use of a different communication style, (e.g.: neutral description Vs dramatization), that requires the story to be engaging. In ACube many Ahab's Legs were introduced because abstract requirements, short and not contextualized sentences, were translated into full-fledged narrations, characterized by concrete setting, well defined interacting personas and a plot with a dramatic tension for raising the engagement of stakeholders.

In the example of Figure 1, the PDA is an Ahab's Leg, because the system requirement about alert messages does not detail the kind of devices to use. However, when this requirement is translated into a scenario, the analyst had decide how to explain this interaction to stakeholders, so the PDA was chosen.

Even if requirements can be expressed in a "simple" and abstract way to system designers, they are more difficult to present to stakeholders because they do not share the same technical background. Conversely, narrative or pictorial scenarios are communicative, because they are more concrete. In order to instantiate requirements in a concrete spatial-temporal context and to raise the dramatic tension to the story, brand new details must be invented by the analyst and added to scenarios, for examples actors have names and actions happen in precise places. Ahab's Legs are often unavoidable, because they are an effect of the translation, it is not possible to instantiate a requirement without providing additional details.

Despite the intended effect of Ahab's Legs (i.e. make scenarios concrete), it may be the case that stakeholders focus their attention too much on non-central aspects. Anyway Ahab's Legs do not necessarily represent a problem, unless they divert stakeholders attention from the important aspects of the story. This causes a waste of valuable time during the session, since comments raised on Ahab's Legs do not impact any part of the actual requirements, as it happened with the PDA in our example.

We recognize that details are important to make scenario concrete, it is not realistic to fully remove them before a validation session. Instead, we can keep Ahab's Legs and assess a way to limit their negative impact on time wasted during requirement validation sessions, namely to distract stakeholders from important details.

C. Dealing with Ahab's Legs

As noted also in a previous paper [5], during validation session, it is important to identify all the sources of *distraction* for stakeholders (and Ahab's Leg is a prominent source),

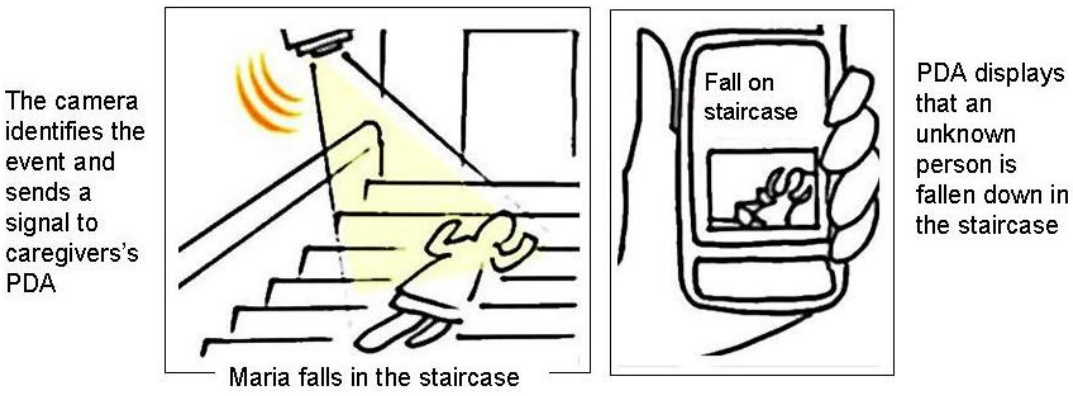

Figure 1. Pictorial representation of a scenario in ACube project.

in order to limit their impact on the session and to maximize the effectiveness of the feedback received. Ahab's Legs are a complex problem and we identify multiple dimensions that deserve further analysis and discussion. However, only one will be selected for experimental validation.

Mandatory Vs optional. An open point concerns the nature of the Ahab's Legs. Some of them are mandatory to the translation since they provide fundamental elements to make the story concrete and believable. Stakeholders need concrete details to sustain imagination and envisage the functionality. It is the case of the PDA as a device to access to a system functionality, in the example of Figure 1. Conversely, other details are optional and choreographic. They are used just to *increase* the stakeholder engagement to the story. An instance of this second kind in the example of Figure 1 is the name of the patient, i.e. Maria. Different kind of Ahab's Legs may have a different impact on distracting stakeholders, as the discussion is not expected to deviate a lot on details that are clearly choreographic. However, it can be interesting to study intermediate cases and assess if the classification given by the analyst matches the classification that stakeholder with a different background would do.

Cardinality. As some Ahab's Legs are more *optional* than others, they could be easily removed from scenarios, so the analyst could control the amount of Ahab's Legs to adopt. In this case a trade-off need to be explored. Scenarios with many Ahab's Legs are very concrete, but with a high risk of distraction. Whereas scenarios with few Ahab's Legs are very abstract and difficult to present to stakeholders. However, an important factor to consider here is the influence of implicit, contextual and personal background knowledge. In fact, narrative stories (such as scenarios) strongly rely on common experience, knowledge and perception of the *real* world everytime some of the properties of their *artificial* world are not explicitly described. For example, Melville does not specify that all the members of the crew of the Pequod have two legs, he relies on the common knowledge of the *real* world [1], where having two legs is a general property of humans. This fact can limit the possibility to avoid some kind of Ahab's Legs by abstracting the description, since this abstraction can lead to a series of hypotheses from the stakeholders based on their own experience, that can completely corrupt the message in the scenario. In other words, by removing too many Ahab's Legs, a scenario may become so abstract that a stakeholder would mentally add her/his own Ahab's Legs, and the analyst would lose any possibility to control their impact.

Stakeholder awareness. One could think of comparing scenarios with Ahab's Legs and scenarios without them, but this is not realistic. We recognize that details are important to make scenario concrete, it is not realistic to fully remove them before a validation session. Intuitively, stakeholders are able to avoid discussions deemed irrelevant, if the validation session in form of focus group [6] is effectively moderated by a facilitator. The problem is to find ways to highlight irrelevant details but, at the same time, avoiding the risk that stakeholders' attention is attracted by them. In a classical focus group session (also adopted in the ACube validation session), it is up to the facilitator to introduce the objective of the validation and to specify what is the expected kind of discussion that would be useful as a feedback. In practice, the facilitator must tell the stakeholders what are the Ahab's Leg before the validation session begins. The analyst could mention that there are details added just to make the scenarios concrete, but they are not the objective of the discussion.

Exploiting the stackeholder awareness is probably the most plausible approach an analyst would take in order to limit the negative impact of Ahab's Legs. In the next section, we design a controlled experiment devoted to validate if this approach can effectively reduce the distraction during validation sessions. To achieve this objective, we have to set up artificial requirement validation sessions to be conducted with users.

III. Experimental definition and planning

This section describes the definition of the experimental settings in a structured way, following the guidelines of Wohlin et al. [9].

The *goal* of the study is to investigate the requirement validation phase with the *purpose* of evaluating the impact of overspecification details (Ahab's Legs) introduced due to the translation from requirements to scenarios.

Our *quality focus* regards the reduction of distraction when overspecification details are explicitly listed. Being the distraction a mental phenomenon, it is difficult to directly measure it, thus we decided to observe stakeholders' feedbacks and check whether their comments are effective to revise the original requirements. We evaluate the results of this experiment on multiple *perspective*: (1) a researcher interested in studying the impact of overspecifications in requirement validation and (2) a focus group facilitator willing to understand whether to explicitly mention Ahab's Legs when meeting project stakeholders in a requirement validation session.

The *context* of the experiment is composed by *subjects* impersonating the stakeholders in requirement validation sessions and the *objects*, the software systems whose requirements are validated.

A. Context description

This experiment consists of a study with computer science master students. Such subjects are supposed to have already attended general courses on *Software Engineering*, *Software Engineering Project* and *Programming Project*, but also specific courses on requirements such as *Requirements and Design of Software Systems*. In such courses the students work on some software projects that require to discuss and validate requirements. Some subjects may also hold industrial experience as full-time software developers.

The software projects to be used for requirements validations are *MyBanking* and *MyShopping*. *MyBanking* is a home banking application for mobile devices (phones/PDAs), designed to replace credit cards and cash in daily money transactions. It supports money transfer between mobile devices using gestures or from/to contacts in the address book. *MyShopping* is an augmented reality application for mobile devices that displays information on items pointed by the camera. This application is meant to provide nutritional and commercial data about goods in the grocery. Both these applications are chosen keeping in mind a reasonable subjects background, in order to maximize the effectiveness of feedbacks. The two applications come from advanced technological scenarios insight for the near future. They aggregate existing technologies (mobile devices) on which subjects are supposed to already own experience, in order to change current practice in daily contexts (restaurant, grocery).

We defined four scenarios for each application. In order to define them, we simulated the requirement elicitation process by taking advantage of internet researches and technological surveys in literature. We firstly collected a list of requirements for our applications, in form of sentences that describe user-system interactions. Then we identified four macro-functionalities to validate with stakeholders, thus we instantiated requirements into scenarios in order to illustrate these functionalities. Narrative scenarios tell concrete and engaging user-stories of the insight product.

B. Hypotheses formulation and variable selection

Based on the study definition described above, we can formulate the subsequent null and alternative hypotheses to be tested:

- H_0 explicitly mentioning what are the overspecified details (Ahab's Legs) in application scenarios does not significantly reduce the distraction in a requirement validation session.
- H_a explicitly mentioning what are the overspecified details (Ahab's Legs) in application scenarios significantly reduces the distraction in a requirement validation session.

The null hypothesis is *one-tailed*, because according to our experience we expect that when stakeholders are aware that those details have been added just to make scenarios concrete, feedback should be more focused on relevant aspects. Even if explicitly mentioning some details could potentially attract (even unconsciously) stakeholder attention and drive the discussion around them.

The experimental hypotheses suggest that the observable outcome of the experiment, namely the dependent variable, is the *distaction* occurred during the requirement validation session. To observe the distraction, we ask subjects to comment on and give feedback related to some scenarios of the object applications.

Feedback is collected by asking subjects what is the first comment/question that they would ask to the analyst who just presented such a scenario. On each application, for the first two scenarios they are asked to select from a closed list of topics, for the other two scenarios subjects freely formulate questions. Distraction is measured by evaluating if a formulated question/comment involves a revision of the application requirements. In case the comment to the i-th scenario involves features not part of requirements, we classify the i-th feedback as distracted by useless details ($distr_i = 1$). Conversely, if the formulated comment/question is about an aspect that is part of requirements, this is very valuable for the analyst, because it involves a revision of such requirement, so overspecification details have not been distracting ($distr_i = 0$). In case no comment/question is formulated, we assume that the stakeholder fully agrees with the scenario. This is also a no distraction case, as long

as the scenario is clear to the stakeholder. We check clarity with a post questionnaire.

We finally measure the distraction for each subject on the whole session as the sum of the distraction observed on the four scenarios:

$$Distr = \sum_{i \in Scenarios} distr_i \quad (1)$$

In other words, the distraction during a validation session is the number of scenarios where we observed distraction.

The independent variable (the main factor of the experiment) is the presence of the treatment during the requirement validation sessions. The two alternative treatments are (1) the presence or (2) the absence of an explicit mention of the overspecified details (Ahab's Legs) added to the scenarios in order to translate requirements for the validation session.

The treatment consists in appending to the description of the object application an extra paragraph with *some* examples of those Ahab's Legs that we added to scenarios. Without using the term *"Ahab's Leg"*, we just mention that scenarios contain details needed to make them concrete, but that do not come directly from requirements. We also specify that the validation is supposed to discuss important aspects of the application. Our objective is not to make subjects aware of *all* the Ahab's Legs, such that subjects would have to skip them when formulating comments. Our purpose is to explain to stakeholders that details can be more or less relevant for the validation, and they should judge relevance when formulating questions. We give just some example of the less relevant details, the Ahab's Legs, to give a starting idea.

Table I shows the description of *myBanking*, that precedes scenarios. In case of treatment, also the extra paragraph is given (Table I, bottom), but not all the Ahab's Legs are listed.

While we design the experiment in order to study the impact of the Ahab's Legs (independent variable) on distraction during requirement validation (dependent variable), other co-factors that we can not control during the experiment may influence the outcome of the experiment itself. Among co-factors that can potentially influence the dependent variable we post-measure:

- *The lab*: subjects could spend some effort to familiarize with the procedure to comment scenarios, during the first validation session. So we measure if any *learning* effect occurs between the two labs.
- *The system*: since we use two systems, subjects could show different performances on different systems. So the system is also a co-factor.
- The ability of subjects in participating proactively in requirement validation sessions could depend on their background. We hence collect the following relevant information about subjects by using a profiling questionnaire:
  - *Merit*: academic merit may have an influence on the ability to discuss requirements. We classify subjects with an average score of at least 28/30 as *high* merit and the others as *low* merit;
  - *Background*: in the study plan we identify four exams that are relevant for the experiment tasks on requirement validation (e.g., *Requirements and Design of Software Systems*), but some of them are optional. We consider the background of subjects as *full* when all the relevant exams have been passed. Otherwise the background in just *partial*;
  - *Requirement experience*: when a subject already discussed requirements for big/real software projects, her/his experience on requirement validation is considered *high*, *low* otherwise;
  - *Development experience*: whether subjects have previous experience as full-time software developers in industry, their development experience is classified as *high*, *low* otherwise;

For each co-factor, we test if there is any effect on the distraction and if the co-factor interacts with the main factor (explicit mention of Ahab's Legs) to influence the distraction. So, for each co-factor, we can formulate the subsequent null hypothesis:

H_{0c_i} The *co-factor* c_i does not significantly interact with the presence of explicit mention of overspecification details to influence the distraction during the requirement validation session.

C. Experimental design

We adopt a balanced design with two experimental sessions (called *lab 1* and *lab 2*), taking in total one class (up to two hours). Subjects are randomly divided in four groups. The design ensures that each group works on the two applications (*MyBanking* and *MyShopping*), with the two treatments (explicit or non-explicit Ahab's Legs) in all the possible permutations, as shown in Table II. This design allows us to analyze not just the effect of the main factor (treatment), but also the effect of each co-factor using proper statistical test (see Wohlin et al. [9] for a quick introduction).

D. Experimental procedure and material

Before the experiment, subjects attend a seminar on requirement validation, to recall the concepts acquired in past courses. Then, they are involved in a training session very similar to the activity that they are supposed to conduct during the experiment. An example of music sharing online application is presented, together with some scenarios. Subjects impersonate potential users of this system in a requirement validation session. They are guided in the process of formulating questions and comments on scenarios by means of a live discussion. This makes us confident that subjects are aware of the settings of the experiment and of the role of their feedback to scenarios during the

Description myBanking application will change your current thinking about money and payments, and will free you from carrying cash and credit cards. A mobile application, following you everywhere, that eases 1) any money transfer to friends, family members or everyone you desire, 2) payments from everywhere a credit card is accepted, 3) to always keep under control you financial record.
Validation Activity Preamble The aim of a validation session is to discuss with you, stakeholders, about the system functionalities. Your feedback will be useful for allowing the analysts in improving the final product.
Mention to Ahab's Legs It is worth noting, the scenarios, we wrote, contain many details in order to make the story concrete and more engaging to the end-users; these details have no impact on system requirements. Examples of these details are: the place in which the story is set (the walk outside, the fast food), the characters that are involved (Michael, Rose), the brands of mobile device they use, the amount of money (100), the procedure for activating functions (voice, movements), among the others.

Table I
EXAMPLE OF MENTION OF AHAB'S LEGS FOR THE TREATMENT GROUP.

	Group1	Group2	Group3	Group4
Lab 1	MyBanking AL	MyBanking NO	MyShopping AL	MyShopping NO
Lab 2	MyShopping NO	MyShopping AL	MyBanking NO	MyBanking AL

Table II
EXPERIMENTAL DESIGN. AL = EXPLICIT MENTION OF OVERSPECIFIED DETAILS (AHAB'S LEGS), NO = NO EXPLICIT MENTION OF THEM.

experiment. We have to make sure that distraction observed when collecting feedback during the experiment is due to how scenarios are formulated and not by problems in understanding what kind of feedback subjects are supposed to provide.

Before the experiment, we present to the subject the experimental procedure, without making explicit mention of the study hypotheses. We distribute the subsequent material on paper sheets:

1) The profiling questionnaire to fill in order to profile the subjects' experience and background;
2) A description of the *first* application either (1) with or (2) without an explicit mention of what are the overspecified details (Ahab's Legs).
3) Four scenarios to comment, the first and the second scenarios requires to select the topic of the feedback from a list of topics, the third and fourth scenarios allow free feedback.
4) A description of the *second* application either (1) with or (2) without an explicit mention of what are the overspecified details (Ahab's Legs).
5) Four scenarios to comment, the first and the second scenarios requires to select the topic of the feedback from a list of topics, the third and fourth scenarios allow free feedback.
6) A post experiment questionnaire to fill.

During the experiment, the teaching assistant is present to prevent collaboration among subjects and to make sure that the experimental procedure is followed. In particular he checks that scenarios and applications are addressed in the expected order. In order to limit the effect of evaluation apprehension, the experiment is conducted anonymously, but we can correlate profiling answers, feedback to scenarios and post questionnaire for the same subject, because the sheets are fastened by a binder.

The post questionnaire is meant to gain insight about the subjects' behavior during the experiment and to find justification for the quantitative observations. The questionnaire contains 6 questions expressed in the Likert scale [8], from "Strongly agree" (2) to "Strongly disagree" (-2). Questions are about:

- $\mathbf{Q}_1$: Whether the time given to complete the task was enough;
- $\mathbf{Q}_2$: Clarity of the tasks;
- $\mathbf{Q}_3$: Clarity of the descriptions of the scenarios;
- $\mathbf{Q}_4$: Clarity of the application domain;
- $\mathbf{Q}_5$: Whether topics in the lists from which to chose (scenarios 1 and 2) match the subject intended feedback; and
- $\mathbf{Q}_6$: Whether subjects perceived the list of Ahab's Legs as helpful to focus on other relevant details.

E. Analysis method

In order to test the hypothesis on the distraction during requirement validation sessions (H_0), we use a non-parametric test, because it does not impose any constraint on the normal distribution of the population. Moreover, since we collect two measurements for each subject (i.e., with presence or absence of explicit mention of Ahab's Legs), data are intrinsically paired, so we used a paired statistical test, the Wilcoxon one-tailed test. Such a test allows to check whether differences exhibited by the same subjects with different treatments over the two labs are significant. In case the test returns a p-value$<$0.05, we can reject the null hypothesis H_0 and formulate the alternative hypothesis H_a, i.e. the act of mentioning Ahab's Legs significantly reduces the distraction during requirement validation.

While this analysis allows to state that the difference is statistically significant, it is also of practical interest to estimate the magnitude of such difference. For this objective we use the Cohen *d* effect size, as it indicates the magnitude of the main factor effect on the dependent variable. For dependent samples (in context of paired analysis) the effect size is computed as the difference between the means (M_{NO} and M_{AL}) divided by the standard deviation of the (paired) difference between samples (σ_D):

$$d = \frac{M_{NO} - M_{AL}}{\sigma_D} \quad (2)$$

The effect size is considered small for $0.2 \leq d < 0.5$, medium for $0.5 \leq d < 0.8$ and large for $d \geq 0.8$.

The analysis of co-factors, i.e., the test of hypotheses H_{0_i}, is performed using a two-way Analysis of Variance (Anova) and, when present, interactions are visualized using interaction plots. Although Anova is a parametric test, it is considered quite robust also for non-normal and non-interval scale variables.

Regarding the analysis of survey questionnaires, we evaluate each question by verifying that the answers is either "Strongly agree" (2) or "Agree" (1). We test medians, using a one-tailed Mann-Whitney test for the null hypothesis $\widetilde{Q_i} \leq 0$, where 0 corresponds to "Undecided", and $\widetilde{Q_i}$ is the median for question Q_i.

IV. Related works

Many experiments reported in requirements literature focus on the analysis of two or more methodologies, to compare their communication qualities, such as comprehensibility of business requirements, or the complexity and efficacy of their visual notations, or the capacity to let analysts reason about the quality of a set of requirements. For example, in [4] the problem of comparing the communication qualities of two requirements methodologies and notations has been investigated. In particular, a controlled experiment is described aiming at comparing two requirements modeling methods: Use Cases, a scenario-based requirements specification approach exploiting both diagrammatic notations and text, and Tropos, a goal-oriented approach based on a complete diagrammatic notation. The objective was to evaluate the comprehension of requirements models expressed in each method notation. The result shown that Tropos overcame Use Cases in therm of comprehenisibility but seems to be more time consuming. In the line of the assessments of the quality of the requirements, in [2] is described a controlled experiment for studying the completeness and granularity of functional requirements specifications described in two different approaches: Use Cases and Communication Analysis. The latter is a method for the development and computerisation of enterprise Information Systems that focuses on communicative interactions that occur between the systems and its environment. Communication analysis resulted having a greater quality in terms of completeness and granularity.

Another problem relevant to our work is the ambiguity of requirements expressed in natural language and its nocuous effects on the specification and interpretation. This aspect has been investigated in [10]. Here an approach based on machine learning techniques is described to automatically detect this phenomenon when different readers interpret the requirements. Some experiments are also presented to asses the performances of the methods that seem to be good in terms of precision and recall.

The problem of identifying abstraction in a requirement specification is discussed in [3]. Here the problem is that of identifying terms in the problem domain that (also indirectly) encapsulates the scope of the envisioned system and how and to what extent these terms can be abstracted into more general terms by the requirements engineer when modelling the domain. The paper presents a new abstraction identification technique, namely RAI, and compare its performance with classical techniques based on human judgement.

Finally, the problem of how to describe requirements via scenarios to support their validation with stakeholders is investigated in a previous paper [5]. The problem of Ahab's Leg for requirements is introduced and described with the aim of highlighting its properties, in order to know and (try to) limit drawbacks. The present paper builds on top of the previous work, with the purpose of empirically assessing previous observations.

V. Conclusion

While validating requirements on a real project, we identified the problem called Ahab's Leg dilemma. When translating requirements to adoption scenarios, details (Ahab's Legs) need to be added that make scenarios more concrete and engaging to stakeholder. However, the role of this kind of details is controversial, because on the one hand they allow to reason on a concrete example of the system to be implemented, so practical problems can be discussed with stakeholders. On the other hand, Ahab's Legs are not part of the application requirements, they could distract the attention of stakeholders from more important aspects of scenarios and cause wasting of time as the discussion deviates on them.

We conjecture that the Ahab's Leg dilemma is a general problem, that affects the communication between analysts and stakeholders. In this paper we identify a crucial role of Ahab's Legs and we speculate on the positive and negative impact that they may have or requirement validation sessions. We discuss several possible approaches that an analyst may adopt to limit distraction caused by Ahab's Legs. Among them, we identify the most practical and promising one to be assessed empirically.

We propose a metric to measure distraction during requirement validation sessions and we present the design of a controlled experiment devoted to evaluate the impact Ahab's Legs on distraction. In particular, we intend to ask some subjects to comment scenarios when the issue of Ahab's Legs is explicitly mentioned or not mentioned at all.

We plan to conduct this experiment first of all with master students from different universities. Then, in order to study how distraction changes when the background of subjects also changes, we intend to replicate the study with totally different people (e.g., requirement engineers, domain experts and actual stakeholders). Two additional interesting aspects that we intend to investigate in further replications concern the way scenarios are presented (e.g., textual or pictorial representation) and the kind of Ahab's Legs (e.g., simple list or structured ontology). We believe, in fact, that different representations can potentially led to a different impact.

REFERENCES

[1] U. Eco. *Kant and the platypus: essays on language and cognition*. Harvest Books, 2000. (Translation from italian of the book *Kant e l'Ornitorinco, Bompiani, 1997*).

[2] S. España, N. Condori-Fernández, A. González, and O. Pastor. Evaluating the completeness and granularity of functional requirements specifications: A controlled experiment. In *RE*, pages 161–170. IEEE Computer Society, 2009.

[3] R. Gacitua, P. Sawyer, and V. Gervasi. On the effectiveness of abstraction identification in requirements engineering. In *RE*, pages 5–14, 2010.

[4] I. Hadar, T. Kuflik, A. Perini, I. Reinhartz-Berger, F. Ricca, and A. Susi. An empirical study of requirements model understanding: *Use Case* vs. *tropos* models. In S. Y. Shin, S. Ossowski, M. Schumacher, M. J. Palakal, and C.-C. Hung, editors, *SAC*, pages 2324–2329. ACM, 2010.

[5] C. Leonardi, L. Sabatucci, A. Susi, and M. Zancanaro. Ahab's leg: Exploring the issues of communicating semi-formal requirements to the final users. In *CAiSE*, pages 455–469, 2010.

[6] R. K. Merton. *Focused Interview*. Free Press; 2 edition, 1990.

[7] B. Nuseibeh and S. Easterbrook. Requirements engineering: a roadmap. In *Proceedings of the Conference on The Future of Software Engineering*, ICSE '00, pages 35–46, New York, NY, USA, 2000. ACM.

[8] A. N. Oppenheim. *Questionnaire Design, Interviewing and Attitude Measurement*. Pinter, London, 1992.

[9] C. Wohlin, P. Runeson, M. Höst, M. Ohlsson, B. Regnell, and A. Wesslén. *Experimentation in Software Engineering - An Introduction*. Kluwer Academic Publishers, 2000.

[10] H. Yang, A. N. D. Roeck, V. Gervasi, A. Willis, and B. Nuseibeh. Extending nocuous ambiguity analysis for anaphora in natural language requirements. In *RE*, pages 25–34, 2010.