

Exploring how experienced and unexperienced designers use and evaluate a usability-oriented interaction and navigation model

Anna Beatriz Marques
*USES Research Group, Institute of
 Computing
 Federal University of Amazonas
 Manaus, Brazil*
 anna.beatriz@icompu.ufam.edu.br

Simone Diniz Junqueira Barbosa
*Informatics Department
 PUC-Rio
 Rio de Janeiro, Brazil*
 simone@inf.puc-rio.br

Tayana Uchôa Conte
*USES Research Group, Institute of
 Computing
 Federal University of Amazonas
 Manaus, Brazil*
 tayana@icompu.ufam.edu.br

Abstract— Representing quality aspects in models used in the design of interactive systems can support to design solutions with higher quality of use. However, the quality of the designed solutions can be influenced by the designers' experience and by the models' expressiveness for representing the quality aspects. Recently, we proposed USINN (USability-oriented INteraction and Navigation model) to express usability mechanisms in interaction and navigation modeling solutions. In this paper, we present an experimental study conducted with students, characterized as experienced and unexperienced designers, in order to investigate how is USINN adopted and evaluated by designers with different levels of experience in the software industry. The results indicated that the quality of artifacts produced by experienced and unexperienced designers was similar. However, the unexperienced designers indicated higher intention to adopt USINN in the software industry.

Keywords— *interaction model, navigation model, interactive systems usability, usability-oriented design*

I. INTRODUCTION

The software industry has increasingly focused on offering high-quality interactive system to users [1]. However, interactive systems often contain design defects that may cause users problems [2]. Since the system is only perceived by users through what they can see and with what they can interact, the success of an interactive system is inexorably related to the quality of its user interface [3]. In order to improve the quality of user interfaces, we can focus on first designing interaction rather than interfaces [4]. Thus, it is important that the artifacts employed in interaction design allow expressing quality criteria to be met by the interaction.

Usability is a quality attribute that affects the user interaction with the system, when the system provides commands to allow users to undo actions, to validate user requests and to provide appropriate feedback [5]. Representing functional usability features [5] in models used in the design phase is a proposed strategy for use cases, classes and sequence diagrams of UML (Unified Modeling Language) models [6], and MDD (Model-Driven Development) approaches [7].

We believe that adopting a similar strategy for models used in interaction design will allow increasing the usability expressiveness of the models and, consequently, the usability of the solutions designed. As usability expressiveness, we consider the model's capability of representing usability aspects related to its modeling perspective. When using models with high usability expressiveness, the interaction designer can design her/his solutions by more directly reflecting on usability mechanisms to be integrated in the interaction. This strategy can reduce the rework needed to consider usability in later stages of systems engineering [7].

Recently, we defined a notation called USINN (USability-oriented INteraction and Navigation model) to represent usability mechanisms related to the interaction and navigation of interactive systems [8]. Our goal is that designers adopt USINN during the interaction and interface design activities.

We conducted an empirical study in order to investigate how experienced and unexperienced designers use and evaluate USINN as: (i) a notation for modeling interaction and navigation considering usability mechanisms; and (ii) a basis for creating user interface mockups. We analyzed their perception about USINN and the quality of the artifacts they created with USINN.

The following sections describe the background knowledge for our work, the USINN model, and the empirical study that we conducted. Then, we present the general results, the quality metrics considering the industry experience of the participants and the participants' perception of USINN. Finally, we present our concluding remarks and future work.

II. BACKGROUND AND RELATED WORK

A structured representation of user goals and tasks is a key activity in the user-centered design of interactive systems [9]. Although task models are an established approach to this end, representing alternative paths to reach the same goal requires the development of different task models and it is difficult to take advantage of rich control flows and contexts [10]. Such limitations can be addressed by the use of interaction models [10][11].

We consider as interaction model any model that describes the communication between the user and the system, specifying when the user can perform specific tasks to achieve certain goals, when the user can inform input data, and when the system can process the user information and show appropriate content and feedback [11]. Interaction and navigation are interrelated aspects of design, because the user-system interaction can lead to navigation flows [8]. Modeling the navigational perspective according to the way in which the user wishes to explore the application helps to obtain more usable navigational paths [14].

Navigation models comprise basically navigation nodes (a set of information or features that will be presented to users) and navigation flows among nodes [12]. As discussed in [8], USINN was the first model to explicitly integrate interaction and navigation representations. When designing USINN, we kept in mind that models that represent different views of the same solution must be consistent with each other [13].

Furthermore, representing usability aspects in design models is a proposed strategy for use cases, classes and sequence diagrams of UML (Unified Modeling Language) [6] and MDD (Model-Driven Development) approaches [7], but USINN was the only similar approach for interaction models. We believe that increasing the usability expressiveness of the interaction models will also improve the usability of the solutions designed. Using models with better usability expressiveness could enable the designer to reflect on usability aspects to be considered in the interaction.

In order to consider usability aspects in USINN, we drew from the usability mechanisms defined by Juristo et al. [15]. Juristo et al. [5] identified a set of functional usability features (named FUFs) that influence the design of the application, specifically the user interaction. They detailed each FUF in terms of usability mechanisms (which are different ways of designing each FUF) and proposed guidelines to elicit usability requirements related to usability mechanisms [15].

We have also identified a proposal to include usability requirements in navigation models [12]. However, the authors do not provide specific guidelines on how to elicit usability requirements related to navigation (i.e. it is not directed to FUFs). When adopting USINN, one can use the available guidelines to elicit usability requirements to be considered in the solution design. Beyond that, there are evidences regarding the positive effect of the usability mechanisms on development time [5] and final product quality [15][7].

III. USINN: USABILITY-ORIENTED INTERACTION AND NAVIGATION MODEL

The USability-oriented INteraction and Navigation model (USINN) aims to represent both interaction and navigation aspects of an interactive system. At the same time, USINN allows to expand the interaction and navigation elements related to usability. In order to define the necessary usability elements, USINN elements were also based on the usability mechanisms related to the FUFs [15]. Fig. 1 presents the elements of the USINN notation.

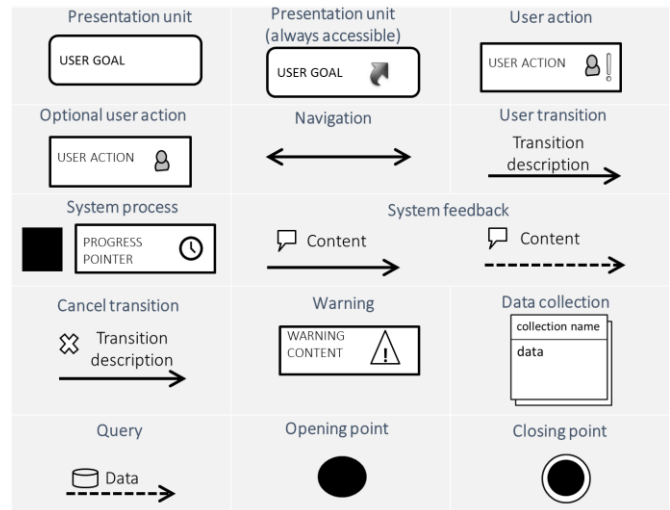


Fig. 1. Summary of the USINN elements representation.

In this section, we provide a brief description of how USINN can visually represent the usability mechanisms associated to the FUFs. Note that, in addition to representing those usability mechanisms, USINN can also represent both interaction and navigation aspects, even if usability requirements are not explicitly represented. The FUFs considered are the following:

Step-by-Step execution: it allows tasks with multiple steps to be represented as a series of navigable windows. Certain tasks require a series of inputs from the user that may not be feasible to perform in a single window. The presentation units and navigation elements (Fig. 1) can be used to represent step-by-step execution. The presentation unit can represent the steps and navigation provides next/previous commands to user.

System Status Feedback: it provides the user with information on the different statuses the system may be in at any given time. The system feedback elements can be used to inform users about the internal status of the system.

Progress feedback: it aims at providing the user with accurate visual feedback on the progress of the current task. The system process with the progress pointer can be used to inform users that the system is processing an action that will take some time to complete. In such cases, the user will need to be able to cancel the task (see abort operation).

Undo: it provides a way for the user to revert the effects of a previously executed action or series of actions within an application. The cancel transition can be used to represent the possibility of undoing user actions.

Abort operation: it provides the means to cancel an ongoing task or to allow for exiting the application altogether. The cancel transition can also be used to represent the possibility of cancelling an ongoing operation. The navigation element enables the user to go back to a particular state in a navigation sequence.

Warning: it entails providing different alert types upon execution of sensitive actions. Certain tasks have potentially serious consequences (may not be reverted, for example). The system may need to verify whether the user actually aims to

proceed with the task to prevent the user from performing tasks by mistake. The warning element can be used to inform users of any action with important consequences. Some warnings may require providing a cancel option to the user.

Multi-level help: it allows the user to access textual help features in different levels of detail throughout the system. The help content can be represented by data collection. An optional user action can be used to represent that the user can view the help content during the interaction.

Preferences: it provides users with a centralized place where they can alter the application settings. The presentation unit can be used to represent a place where the user can configure his/her preferences. The user actions, system process, system feedback and user transitions can detail how the user can configure his/her preferences. The user preferences options can be represented by data collection.

Favorites: it allows the user to bookmark and keep a collection of favorite places within an application. The user action can be used to represent the possibility of bookmarking places or objects in the system. The user favorites can be represented by data collection.

Personal Object Space: it covers the users' needs to arrange and manipulate objects graphically on screen. Similar to the favorites mechanism, user actions and data collection can be used to represent it.

Commands Aggregation: it allows the user to aggregate commands into macro-like structures for ease of batch execution. The data collection can be used to record the sequences of actions the user may perform repetitively and to save them for future use. The user action can represent the possibility for the user to play a previously recorded macro.

We did not consider structured text entry mechanisms, because interaction and navigation models do not consider the concrete aspects of the user interface that are necessary to represent these mechanisms, such as a mask in a text field [6].

Fig. 2 illustrates an example of a USINN diagram that combines the elements described in Fig. 1 to partially represent the interaction of a user with a knowledge management system. The diagram focuses on system status feedback, progress feedback, undo and cancel.

In the knowledge management system illustrated in Fig. 2, the user interaction begins when the user provides his/her data for accessing the system. The system validates the data by consulting the *researcher* data collection. Then, the system provides a suitable feedback for the user about the operation (system status feedback mechanism). If the data is valid, the user can view a knowledge assets list. The user can also filter the list by informing a knowledge asset title. This operation can be undone by the user (undo mechanism).

At any moment during the interaction, the user can create a new knowledge asset. While the new knowledge asset is creating, the system provides the progress feedback about the operation (progress feedback mechanism) and an option for the user to cancel the operation (cancel mechanism). After the operation is concluded, the user can view an updated list of knowledge assets.

The system navigation can be represented by presentation units and navigation flows without representing the user-system interaction. The user-system interaction elements can increment a previously created system navigation model or a single presentation unit. In order to do this, we can add user actions, transitions, warnings, system process, system feedback, data collections, and other USINN elements.

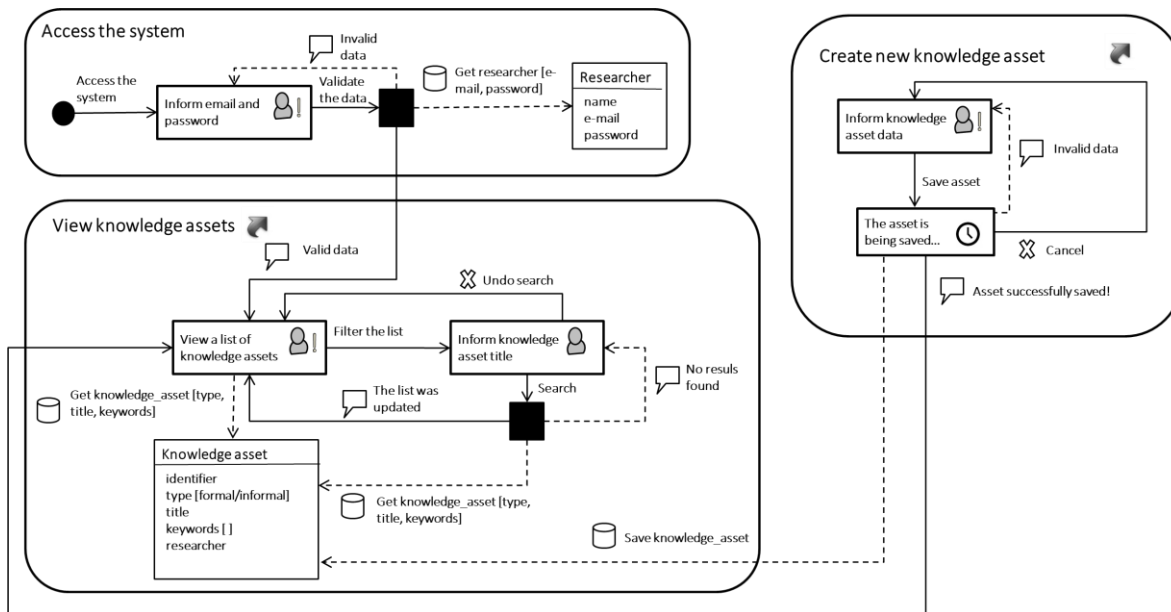


Fig. 2. USINN diagram example.

IV. EMPIRICAL STUDY

To investigate how well USINN supports interaction and interface design, we conducted an empirical study in two stages. In the first stage of the experiment, participants used USINN to model the interaction and navigation of a system. In the second stage of the experiment, the participants used a USINN model as the basis for creating mockups.

A. Goal

The goal of this study was to characterize the use of USINN as a notation for usability-oriented interaction and navigation modeling and as the basis for creating mockups with regards to its ease of use, usefulness, intention to use and the quality of the elaborated artifacts.

B. Participants

In order to evaluate and refine USINN before adopting it in industry, we conducted the experiment in an academic environment in a 5th semester class of a Technology in System Analysis and Development course. We then replicated the experiment in two other classes: (i) a 5th semester class of a Computer Science course and (ii) a 3rd semester class of the Technology in System Analysis and Development course. All classes were related to Human-Computer Interface Design. In total, 55 students participated in the experiment. From those, 11 students had experience in the software industry. We characterized the participants according to their experience with models employed in the analysis and design phases of the software development process, as well as their experience with interface design, in both academia and industry. The student-designers without industry experience had more homogeneous degrees of experience in the design activities and models, related to the conduction of academic projects. Thus, TABLE I presents the different levels of experience of only those designers with experience in industry.

C. Tasks

In the experiment, we asked participants to perform two tasks: (i) to create a model using the USINN notation based on

a scenario and a set of usability requirements; and (ii) to create user interface mockups based on a USINN model.

TABLE I. CHARACTERIZATION OF DESIGNERS' YEARS OF EXPERIENCE.

Participant	Experience using models in the software industry	Experience designing user interface in the software industry
P3	L	L
P7	L	M
P18	M	M
P19	M	M
P27	H	M
P28	M	L
P29	L	L
P37	L	L
P45	M	L
P47	M	M
P52	M	M

Footnote: L=low (up to 1 year); M=medium (between 1 and 2 years); H=high (more than 2 years).

D. Scenarios for modeling task

The scenarios and requirements used in the modeling tasks were related to a web system to manage to-do lists [16]. We described usability mechanisms [15] as functional requirements in accordance to previous studies conducted by Juristo et al. [15] and Carvajal et al. [6]. We aimed to verify whether the participants would be able to model them with the USINN notation. However, describing all usability mechanisms in a same scenario would make the scenario too complex for the experiment. Thus, we decided to create two scenarios about the same system, describing functional requirements associated with different usability mechanisms, as summarized in TABLE II.

TABLE II. FUNCTIONAL REQUIREMENTS AND USABILITY MECHANISMS DESCRIBED IN THE SCENARIOS FOR THE MODELING TASK.

	Requirement	Usability mechanism
Set of functional requirements 1	R01. To create tasks: In order to create a new task, the user has to select the type of the task and to inform the name, the description, the task due date and the estimated time for concluding the task. The system must provide the status feedback until the task is created.	System status feedback
	R02. To edit tasks: The user can update the due date and the estimated time for concluding the task. When the user requires the data update, the system must alert that the changes cannot be undone.	Warning
	R03. To delete tasks: The user can delete tasks. When the user deletes a task, the system must ask the user whether he/she wants to undo the operation.	Undo
	R04. To list tasks with tips about task due to the current date: For each task, the system must present a tip about the due date of the task.	Multilevel help
	R05. To bookmark tasks as favorites: The system must enable the user to bookmark any task as favorite. Then, the task has to be included in a favorites list.	Favorites
	R06. To create macros: The user can create sequences of actions he/she may perform repetitively and save them for future usage. In order to create a sequence, the user has to inform the actions in the order they have to be performed and a name for the macro.	Commands aggregation
	R07. To execute macros: The system must enable the user to run a macro previously created.	

	Requirement	Usability mechanism
Set of functional requirements 2	R01. User authentication: The user can access the system by informing his/her username and password. While the system is validating the user data, the system must show a message informing “authenticating, please wait”.	Progress feedback
	R02. To create a user: To create a new user, the system must show a form with three steps. In the first step, the user has to inform his/her personal data: name, email, gender, and date of birth. In the second step, the user has to inform his/her phone number, address, city, and country. In the third step, the user has to inform the username, password, password confirmation, and a security question. During this process, the user can navigate for the next or previous step without losing the data previously inputted.	Step-by-step execution
	R03. To define preferences: The system must provide the user with an option to define his/her preferences: maximum number of tasks to be displayed per page and maximum number of tasks to be displayed on a map.	Preferences
	R04. To organize tasks list through maps: A map is a set of objects represented in a canvas for the user manipulates and uses. Each list must be represented on a map containing a title and a set of tasks, according to the user preferences defined in R03. The user can graphically organize the list inside the map and create new maps.	Personal Object Space
	R05. To move task lists among maps: The system must enable the user to move a list of tasks from one map to another. For this, the user has to select the list and to choose the map to which the list will be moved.	
	R06. To create list maps: In order to create a new list map, the user has to inform name for the map.	
	R07. To exit the system: The system must provide an option for the user exit from the system by a single action, i.e., by clicking in an exit option.	Abort operation

E. Models for prototyping tasks

To perform the prototyping task, the participants used a USINN model describing a system to support the knowledge management of a research group (Fig. 2 partially illustrates a USINN model used in this stage). We created two different USINN models that represent functional requirements associated with different usability mechanisms (TABLE III).

TABLE III. FUNCTIONAL REQUIREMENTS AND USABILITY MECHANISMS DESCRIBED IN THE MODELS FOR PROTOTYPING TASKS.

	Requirement	Mechanism
USINN Model 1	R01. To edit researcher data	Step-by-step execution
	R02. To list knowledge items and bookmark items as favorites	Favorites
	R03. To organize items among maps	Personal Object Space
	R04. To create maps of knowledge items	Personal Object Space
	R05. To obtain information about items through tips	Multilevel help
USINN Model 2	R01. To access the system	System status feedback
	R02. To create knowledge items	Progress feedback
	R03. To exit the system	Abort operation and Warning
	R04. To search knowledge items	Undo
	R05. To create and execute macros	Commands aggregation
	R06. To define preferences	Preferences

The requirements were balanced between the scenarios and models in such a way that they presented comparable complexity. We considered (i) the number of requirements to balance the scenarios and (ii) the number of elements to balance the models (we considered user actions, system process, presentation unit, data collection, warning, and system progress). The association among requirements and usability mechanisms was not described in the scenarios. We can view part of the USINN model 2 in Fig. 2, representing the R01, R02 and R04 functional requirements.

F. Metrics

To characterize the USINN model, we used metrics for the quality of the produced artifacts and the perception of the participants. We defined quality metrics based on quality goals drawn from a previous work on conceptual models [17]:

Completeness: it defines how much the artifact presents the necessary information according to its purpose. Omission defects reduce the completeness (see TABLE IV).

Correctness: it defines how much the artifact employs the elements and relationships according to the notation syntax and describes correctly the application domain according to the available information. Inconsistency, incorrect fact and ambiguity defects reduce the correctness (see TABLE IV).

TABLE IV shows the taxonomy proposed by Lopes et al. [1] for classification of defects in interaction models and mockups. Different defects can have different severity levels. The severity level may indicate a higher or lower completeness and correctness of the artifacts.

We defined the following formula to calculate the correctness and completeness from the number of defects identified in an artifact and the severity level (see TABLE V) of the defects:

$$\begin{aligned}
 \text{indicator} &= 1 - \text{defects, where:} \\
 \text{defects} &= \frac{(n_req_low/n_req) + 2*(n_req_medium/n_req) + 3*(n_req_high/ n_req)}{6} \\
 n_req_low &= \text{number of requirements with low severity defects} \\
 n_req_medium &= \text{number of requirements with medium severity defects} \\
 n_req_high &= \text{number of requirements with high severity defects} \\
 n_req &= \text{number of requirements described in an artifact (model or mockup)}
 \end{aligned}$$

The correctness and completeness of an artifact can range from 0 (worst result) to 1 (best result).

TABLE IV. DEFECT TAXONOMY FOR INTERACTION MODELS (IM) AND MOCKUPS (M) DEFINED BY LOPES ET AL. [1].

Defect Types	Art.	Description
Omission	IM	The omission or negligence of any necessary information to solve the problem in the interaction diagram.
	M	The omission or negligence of any information needed for the mockup solution.
Ambiguity	IM	A poor definition of certain information in the interaction diagram, which may lead to multiple interpretations.
	M	A poor definition of certain information in a mockup, which may lead to multiple interpretations.
Incorrect Fact	IM	Misuse of the elements from the interaction diagram for the interpretation of those involved.
	M	Misuse of the interface elements during the mockup development, allowing an incorrect interpretation.
Inconsistency	IM	Conflicting information between the elements of the interaction diagram and the information needed to solve the problem.
	M	Conflicting information between the elements of the mockup and the information needed to solve the problem.

The metrics for participants’ perception were based on the Technology Acceptance Model 3 (TAM 3) [18], which aims to assess a user’s perception about the usefulness and ease of use of a proposed technology. Based on the TAM model, we have the following metrics:

Perceived usefulness: it evaluates the degree to which a person believes that using USINN would enhance his/her job performance.

Perceived ease of use: it evaluates the degree to which a person believes that using USINN would be free of effort.

Self-predicted future usage: it indicates whether users would prefer to use USINN over other models.

TABLE VI presents the statements of our questionnaire based on TAM 3. We used a 5-point ordinal scale, using the

opposing statement question format in order not to influence the participant response (positive or negative).

TABLE V. DEFECT SEVERITY LEVELS.

Severity level	Description
Low	The defect in the model/mockup does not affect the comprehension and understandability of the artifact by the reader.
Medium	The model/mockup is incomplete according to the requirements, is ambiguous in the representation, or used elements of the model notation incorrectly.
High	The omission of one or more requirements in the model/mockup or notation errors affects the understanding of the artifact.

In other words, each question contains two opposite statements that represent the maximum and minimum possible values (5 and 1), and in which the value 3 is considered to be a neutral opinion. Also, we replaced the term “[task]” with the terms “interaction and navigation modeling” or “creating mockups” according to the task the participant was carrying out, generating two different questionnaires that we used in the different stages of the experiment.

G. Artifacts

To support the experiment execution, we produced: (i) a consent form; (ii) a participant characterization form; (iii) two scenarios for the modeling task; (iv) two USINN models for the prototyping task; (v) an evaluation questionnaire about the modeling task; and (vi) an evaluation questionnaire about the prototyping task.

H. Execution

The participants received training on interaction and navigation modeling with USINN. The training occurred during three lessons (each lesson lasted for one hour and forty minutes), which included practical modeling and prototyping activities. Fig. 3 illustrates the experiment execution.

TABLE VI. QUESTIONS OF SUBJECTIVE METRICS QUESTIONNAIRE.

Perceived usefulness of USINN	
Using USINN in my job would enable me to accomplish [task] more quickly.	Using USINN in my job would not enable me to accomplish [task] more quickly.
Using USINN would improve my job performance in [task].	Using USINN would not improve my job performance in [task].
Using USINN to [task] would increase my productivity.	Using USINN to [task] would not increase my productivity.
Using USINN would enhance my effectiveness on [task].	Using USINN would not enhance my effectiveness on [task].
Using USINN would make it easier to do [task].	Using USINN would make it more difficult to do [task].
I would find USINN useful in [task].	I would not find USINN useful in [task].
Perceived ease of use of USINN	
Learning to [task] with USINN would be easy for me.	Learning to [task] with USINN would be difficult for me.
I would find it easy to get USINN for [task].	I would find it difficult to get USINN for [task].
USINN elements would be clear and understandable.	USINN elements would be confused and hard to understand.
It was easy to become skillful using USINN for [task].	It was difficult to become skillful using USINN for [task].
It is easy to remember how to perform [task] using USINN.	It is difficult to remember how to perform [task] using USINN.
I would find USINN easy to use in [task].	I would not find USINN easy to use [task].
Self-predicted future usage of USINN	
Assuming USINN would be available on my job, I predict that I will use it on a regular basis in the future.	Assuming USINN would be available on my job, I predict that I will not use it on a regular basis in the future.
I would prefer using USINN to other models (like UML models) for [task].	I would prefer using other models (like UML models) to USINN for [task].

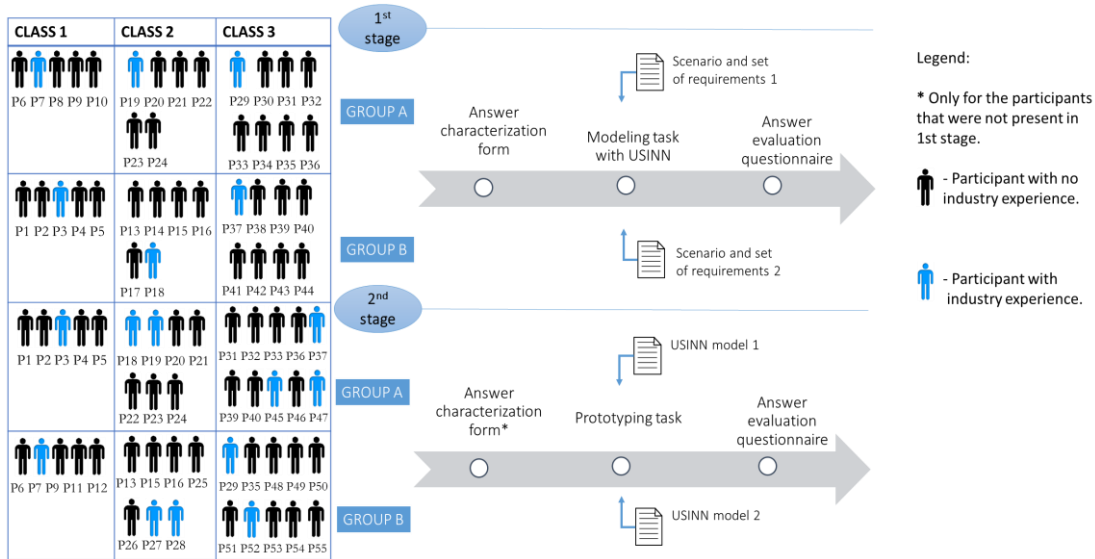


Fig. 3. Execution process of the feasibility study.

The stages of the experiment took place in different days. Some participants were present only in one of the stages of the experiment, while others were present in both stages. For example, the participants P3, P7, P18, P19, P29 and P37 were present in both stages. To enable the use of the different scenarios and models, the participants were organized into two groups (Group A and Group B). To balance the groups, we considered the participants' previous experience in the industry.

In the first stage, the participants elaborated an interaction and navigation model using the USINN notation, based on a scenario and a set of functional requirements (see TABLE II). After concluding the modeling task, the participants received a questionnaire to indicate their perceived usefulness, perceived ease of use and self-predicted future usage of USINN for interaction and navigation modeling.

In the second stage, the participants created mockups based on a USINN interaction and navigation model (see TABLE III and Fig. 2). After concluding the prototyping task, the participants received a second questionnaire to indicate their perceived usefulness, perceived ease of use and self-predicted future usage of USINN as a basis for creating mockups.

V. GENERAL RESULTS OF QUALITY METRICS

We defined correctness and completeness metrics as quality indicators of the produced artifacts. To calculate the quality metrics, we analyzed the artifacts that participants produced (models and mockups) to identify defects that could affect their quality.

To define the artifacts correctness, we performed an inspection to detect inconsistency, incorrect fact, and ambiguity defects. To define the artifacts' completeness, the inspection aimed to detect omission defects. The taxonomy proposed by Lopes et al. [1] guided us during the classification of the defects that were identified during the inspection. We

classified the severity level of the defects, following the rules described in TABLE V. One researcher identified the defects and classified their severity level in the models/mockups and a second researcher validated the results.

A. Models correctness and completeness

Fig. 4 illustrates the model correctness and completeness means per class. We observed means greater than 0.8 for correctness, indicating a positive result for this indicator.

We observed means between 0.5 and 0.7 for completeness, indicating a need for improvement related to this indicator. These results may have two causes: either the participants did not know how to represent some requirements through the USINN elements or they did not remember the model notation, causing them to omit certain elements.

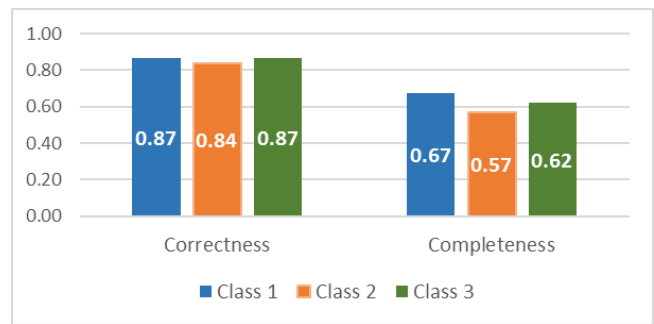


Fig. 4. Models correctness and completeness.

To better understand the quantitative results, we analyzed the comments provided by the participants in our questionnaire. To present the quotes, we have identified the participants by PX-CY, which indicates Participant X from Class Y. Regarding the results of the models correctness and completeness, we highlight the following quotes:

"It is necessary to have domain knowledge about the elements used to create (the model), because we can

understand the requirement, but we can't use the model correctly" – P6-C1

"I need to use USINN more in order to apply it properly. I had difficulties in modeling some requirements" – P3-C2

Since different artifacts were used by groups A and B, we conducted hypothesis tests on the samples to verify whether there was significant difference between the results of the groups. We defined the following null (H_0) and alternative (H_1) hypotheses:

H_{01} : There is no significant difference in the correctness of the models between groups A and B.

H_{11} : There is a difference in the correctness of the models between groups A and B.

H_{02} : There is no significant difference in the completeness of the models between groups A and B.

H_{12} : There is a difference in the completeness of the models between groups A and B.

We verified the normality of the samples by conducting the Shapiro-Wilk test, once the sample sizes were under 50 participants [19]. When we obtain values greater than 0.05, the sample is considered of normal distribution. Conversely, for values less than or equal to 0.05, the sample is not considered normal. For normal distributions, we applied the t-test [20] to test the hypotheses. Otherwise, we applied the nonparametric Mann-Whitney test [19].

The hypotheses testing results are presented in TABLE VII. Values greater than or equal to 0.05 fail to reject the null hypothesis, i.e., they indicate that there was no significant difference between the groups. Conversely, values less than

0.05 reject the null hypothesis and indicate that there was significant difference between the groups.

TABLE VII. RESULTS FROM THE HYPOTHESES TESTS REGARDING THE QUALITY INDICATORS OF THE MODELS.

Metric	Class	Group	Normality Test	Hypothesis Test	
Correctness (H_{01} e H_{11})	Class 1	A	0.700	t Test	0.323
		B	0.464		
	Class 2	A	0.594	t Test	0.956
		B	0.095		
	Class 3	A	0.861	t Test	0.980
		B	0.187		
Completeness (H_{02} e H_{12})	Class 1	A	0.650	t Test	0.507
		B	0.361		
	Class 2	A	0.199	t Test	1.000
		B	0.641		
	Class 3	A	0.008	Mann-Whitney	0.721
		B	0.023		

The hypothesis testing results indicate that none of the null hypotheses were rejected. Thus, the different scenarios did not influence the models' correctness and completeness between the groups. It is important to note that this result was expected because it indicates that the different sets of functional requirements did not bias the results of the experiment.

B. Mockups correctness and completeness

Fig. 6 illustrates the mockups correctness and completeness means per class. We obtained means greater than 0.8 for correctness. On the other hand, we obtained means less than 0.8 for completeness in classes 2 and 3. Fig. 5 shows some examples of mockups created by the participants.

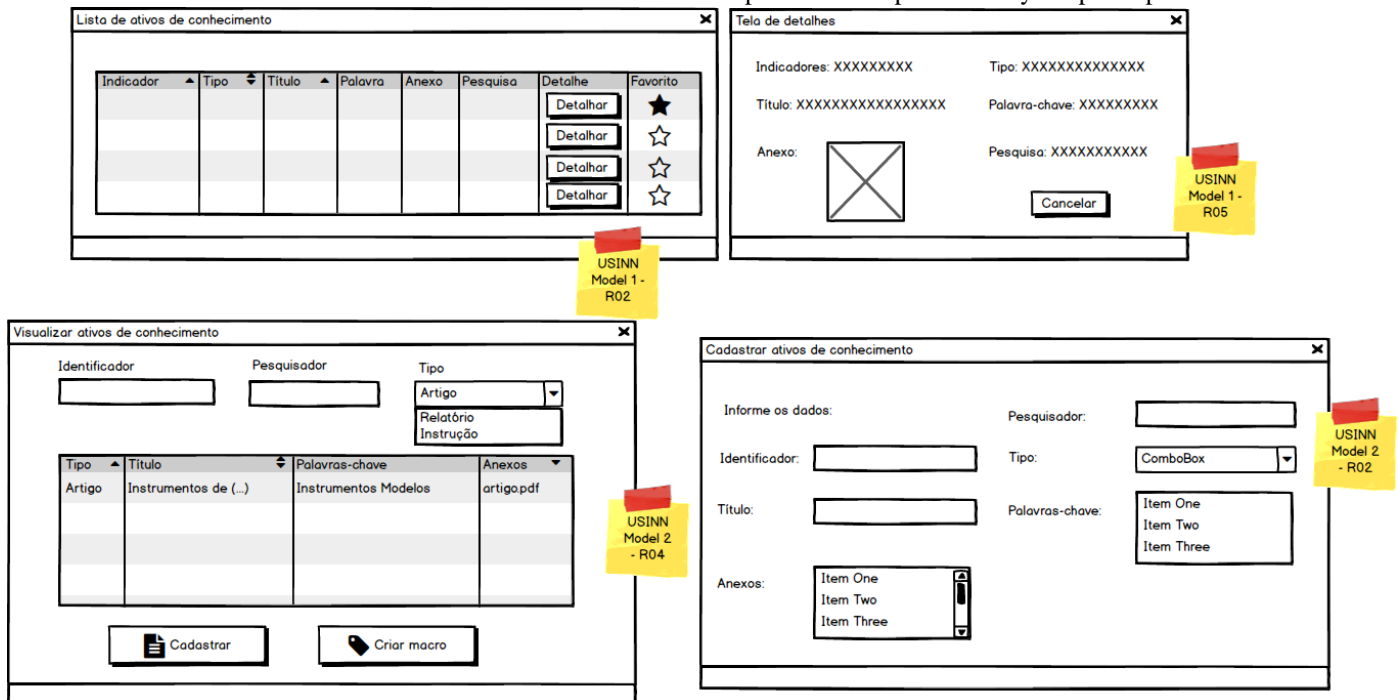


Fig. 5. Examples of mockups created by the participants.

The participants created the mockups using paper and pencil, but in order to provide a better visualization, we transcribed some mockups using Balsamiq Mockups.¹

Some quotes suggest difficulties in understanding the model for creating the mockups:

“The aspects that represent data left me confused. I didn’t know if it was a data view or a table for the user to fill” – P1-C2.

“I believe I have understood what was being proposed, but I still have questions about the interface elements” – P11-C3

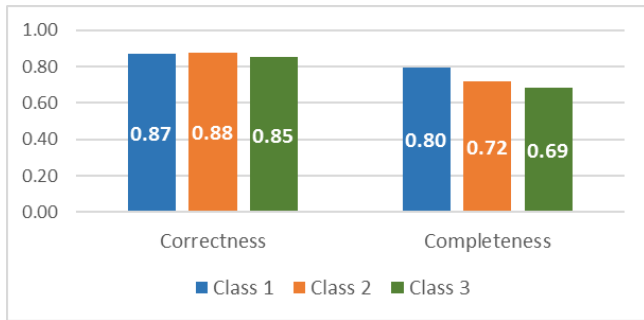


Fig. 6. Mockups correctness and completeness.

We conducted a hypothesis test to verify whether the use of different models between the groups influenced the quality of the mockups produced by the participants (TABLE VIII).

We defined the following null (H_0) and alternative (H_1) hypotheses:

H_{01} : There is no significant difference in the correctness of the mockups between groups A and B.

H_{11} : There is a difference in the correctness of the mockups between groups A and B.

H_{02} : There is no significant difference in the completeness of the mockups between groups A and B.

H_{12} : There is a difference in the completeness of the mockups between groups A and B.

TABLE VIII. RESULTS FROM THE HYPOTHESES TESTS REGARDING THE QUALITY INDICATORS OF THE MOCKUPS.

Metric	Class	Group	Normality Test	Hypothesis Test	
Correctness (H_{01} e H_{11})	Class 1	A	0.008	Mann-Whitney	1.000
		B	0.716		
	Class 2	A	0.009	Mann-Whitney	0.097
		B	0.670		
	Class 3	A	0.544	Mann-Whitney	0.912
		B	0.021		
Completeness (H_{02} e H_{12})	Class 1	A	0.205	Teste t	0.286
		B	0.250		
	Class 2	A	0.484	Teste t	1.000
		B	0.960		
	Class 3	A	0.865	Teste t	0.862
		B	0.264		

¹ <https://balsamiq.com/>

The hypothesis testing results indicate that the null hypotheses were not rejected. Thus, the use of different models did not influence the results. This result is also the expected result, indicating that the different models used did not bias the results of the different groups.

VI. EXPLORING THE QUALITY METRICS CONSIDERING PARTICIPANTS’ EXPERIENCE IN INDUSTRY

In order to investigate whether the designers’ industry experience had an effect on the quality of artifacts they created with USINN, we decided to explore the quality metrics from a different perspective. In this section, we compare the correctness and completeness of the models and mockups created by the experienced and unexperienced designers.

A. Models correctness and completeness

Fig. 7 illustrates the correctness and completeness means comparing the models created by experienced and unexperienced designers. The correctness and completeness of the models created by unexperienced designers were equal to 0.85 and 0.61, respectively. In turn, the correctness and completeness of the models created by experienced designers were equal to 0.89 and 0.65, respectively. We observe that the experienced and unexperienced designers created models with similar quality indicators.

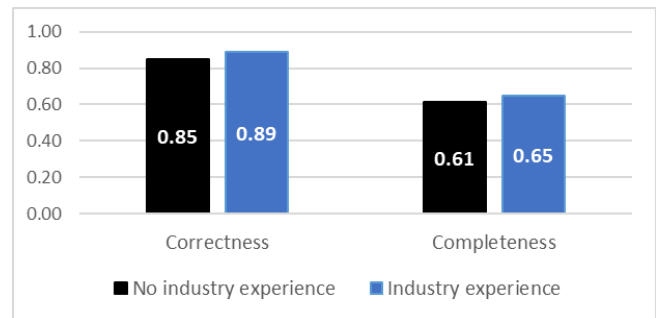


Fig. 7. Models correctness and completeness considering the industry experience.

Fig. 8 illustrates the mean of defects identified in the models per type of defect according to the taxonomy described in Lopes et al. [1]. We observe that the mean of inconsistency defects was less than 1.0 defect/model for both experienced and unexperienced designers. This result indicates that inconsistency defects were identified in few models.

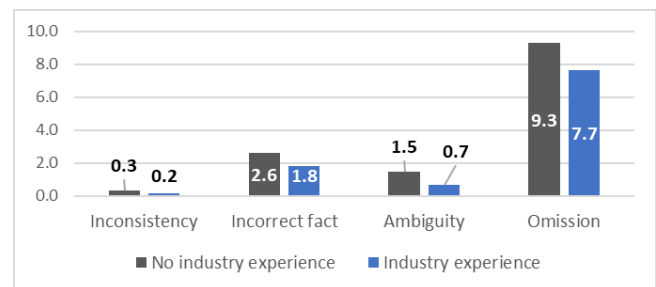


Fig. 8. Mean of the number of defects per type in models.

On the other hand, the means of omission defects were greater than the other types of defects. The mean of omission

defects in the models created by unexperienced designers was equal to 9.3 defects/model. This value was greater than the mean of omission defects in models created by experienced designers, which was equal to 7.7 defects/model.

The unexperienced designers made more mistakes that affect the correctness (2.6 incorrect fact defects per model and 1.5 ambiguity defects per model) when compared to the experienced designers (1.8 incorrect fact defects per model and 0.7 ambiguity defect per model). However, the results indicate that both experienced and unexperienced made more mistakes that affect the completeness than the correctness.

In Fig. 9, we can see the mean of defects identified in the models per severity level. We note that the means of defects with medium severity level were greater than the other severity levels. We can observe that the means of defects with low and medium severity levels were greater in models created by unexperienced designers. The means of defects with high severity levels were similar in models created by experienced and unexperienced designers.

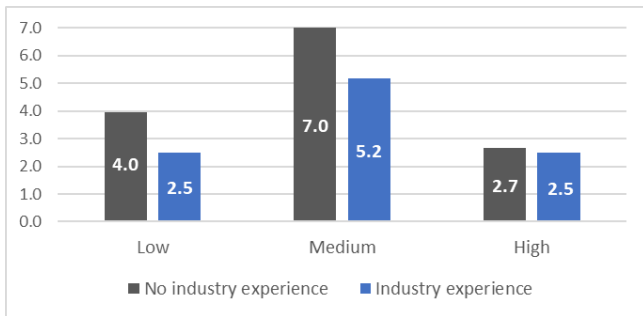


Fig. 9. Mean of the number of defects per severity level in models.

B. Mockups correctness and completeness

Fig. 10 illustrates the correctness and completeness means comparing the mockups created by experienced and unexperienced designers. The correctness of mockups created by experienced designers was equal to 0.89. This value is similar to the correctness of mockups created by unexperienced designers, which is equal to 0.86.

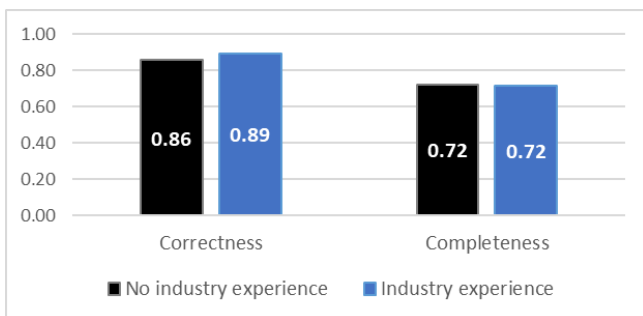


Fig. 10. Mockups correctness and completeness.

When we analyze the mean of defects identified in the mockups per type of defect in Fig. 11, we also note similar values for both experienced and unexperienced designers. We can see that the few incorrect fact defects were identified in the mockups, once the means are less than 1.0 for both experienced and unexperienced designers. We did not observe

ambiguity defects in the mockups. Thus, we can consider that the quality of the mockups created by both experienced and unexperienced designers was similar.

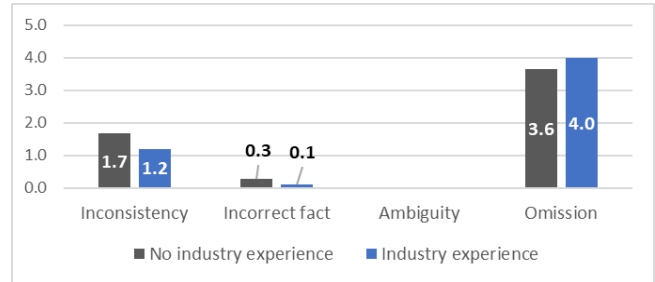


Fig. 11. Mean of the number of defects per type in mockups.

In Fig. 12, we can see the mean of defects identified in the mockups per severity level. We note that there is no defect with low severity in mockups created by experienced designers. We also note that there are few defects with low severity levels in mockups created by unexperienced designers.

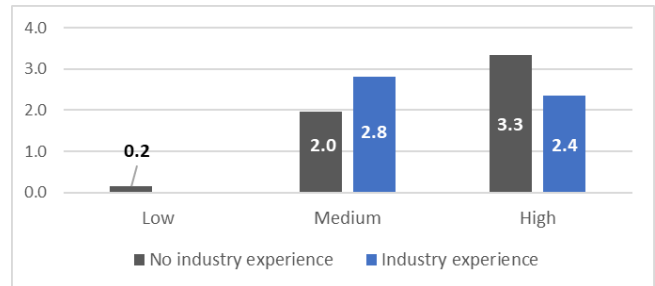


Fig. 12. Mean of the number of defects per severity level in mockups.

We observe the mean of defects with medium severity level in mockups created by experienced designers (2.8 defects per mockup) was greater than that created by unexperienced designers (2.0 defects per mockup). On the other hand, we can note that mean of defects with high severity level in mockups created by experienced designers (2.4 defects per mockup) was less than that created by unexperienced designers (3.3. defects per mockup).

This contradiction can be explained by the fact that when we have an omission defect with high severity level, it indicates that a requirement was completely omitted in the artifact. Thus, possible defects that affect the correctness cannot be identified, once the designer does not model the requirement. However, these defects are reflected in the artifact completeness.

VII. RESULTS OF PARTICIPANTS' PERCEPTION METRICS

As our questionnaires were adaptations of the original statements of the TAM model [18], it is important to evaluate the reliability of our data collection instrument. To assess the reliability of the questionnaire, we conducted a Cronbach's alpha statistical test on the study sample using SPSS.²

² <https://www.ibm.com/analytics/data-science/predictive-analytics/spss-statistical-software>

Regarding the questionnaire that we used to evaluate the modeling stage, we have obtained a reliability value equal to 0.940. When we evaluate the questionnaire that we used in the prototyping stage, we have obtained a reliability value equal to 0.948. To consider a data collection instrument reliable, the test should indicate a reliability level higher than 0.8 [21]. Thus, the results demonstrate that our questionnaire is a reliable instrument based on the results of several empirical studies using the TAM model [22].

TABLE IX shows an overview of the questionnaire results. To quantitatively analyze the participants' answers, we assigned values to an ordinal scale, ranging from 1 (strongly agree to the negative statement) to 5 (strongly agree to the positive statement). To summarize the results based on the ordinal scale, we calculated the median and interval of each item.

Considering that a median equal to 5.0, which is equivalent to "strongly agree to the positive statement", would be the best

possible result, we observed positive perceptions of the experienced designers with regards to the statements (Q1 to Q6) for perceived usefulness of USINN for interaction and navigation modeling. On the other hand, the median indicates a tendency to neutrality in the perception of unexperienced designers regarding the statements Q2 (improve performance) and Q3 (increase productivity). In relation to perceived ease of use of USINN for interaction and navigation modeling, the medians of both experienced and unexperienced designers ranged from 3.5 to 4.5.

With regards to the perceived usefulness of USINN as basis for creating mockups, the medians of the statements were equal to 4 for both experienced and unexperienced designers, except for statement Q5 (makes prototyping easier). Regarding the perceived ease of use of USINN as basis for creating mockups, the medians of the experienced designers' perceptions were equal to 3 for all the statements, while the medians of the unexperienced designers' perceptions were 4 (Q1, Q2, Q3, Q4 and Q6) and 5 (only for Q5 – easy to remember).

TABLE IX. MEDIAN VALUES OF THE PARTICIPANTS PERCEPTION METRICS (ORDINAL SCALE: 5 STRONGLY AGREE TO THE POSITIVE STATEMENT – 1 STRONGLY AGREE TO THE NEGATIVE STATEMENT).

Participants perception metrics	Statements	Experienced designers		Unexperienced designers	
		Median	Interval	Median	Interval
Perceived usefulness of USINN for interaction and navigation modeling	Q1 – model more quickly	4	2 - 5	4.5	2 - 5
	Q2 – improve performance	4	2 - 5	3.5	2 - 5
	Q3 – increase productivity	4	2 - 5	3.5	3 - 5
	Q4 – improve effectiveness	4	2 - 5	4	2 - 5
	Q5 – makes modeling easier	4	2 - 5	4.5	2 - 5
	Q6 - useful	5	3 - 5	4.5	2 - 5
Perceived ease of use of USINN for interaction and navigation modeling	Q1 – easy to learn	3	1 - 5	4	1 - 5
	Q2 – easy to perform	3.5	1 - 5	4	1 - 5
	Q3 – clear and understandable	4	1 - 5	3.5	1 - 5
	Q4 – easy to become skillful	4	2 - 5	4.5	1 - 5
	Q5 – easy to remember	4	1 - 5	4	1 - 4
	Q6 – easy to use	3.5	1 - 5	4	1 - 5
Perceived usefulness of USINN as basis for creating mockups	Q1 – prototype more quickly	4	2 - 5	4	2 - 5
	Q2 – improve performance	4	1 - 5	4	2 - 5
	Q3 – increase productivity	4	2 - 5	4	1 - 5
	Q4 – improve effectiveness	4	1 - 5	4	1 - 5
	Q5 – makes prototyping easier	3	1 - 5	4	3 - 5
	Q6 – useful	4	1 - 5	4	3 - 5
Perceived ease of use of USINN as basis for creating mockups	Q1 – easy to learn	3	1 - 5	4	2 - 5
	Q2 – easy to perform	3	1 - 5	4	1 - 5
	Q3 – clear and understandable	3	1 - 5	4	1 - 5
	Q4 – easy to become skillful	3	1 - 5	4	1 - 5
	Q5 – easy to remember	3	1 - 5	5	1 - 5
	Q6 – easy to use	3	1 - 5	4	3 - 5
Self-predicted future usage of USINN	Q1 – predict future use	4.5	1 - 5	2.5	1 - 5
	Q2 – prefer to use for modeling	3.5	1 - 5	2.5	1 - 5
	Q3 – prefer to use for prototyping	3	1 - 5	3	1 - 5

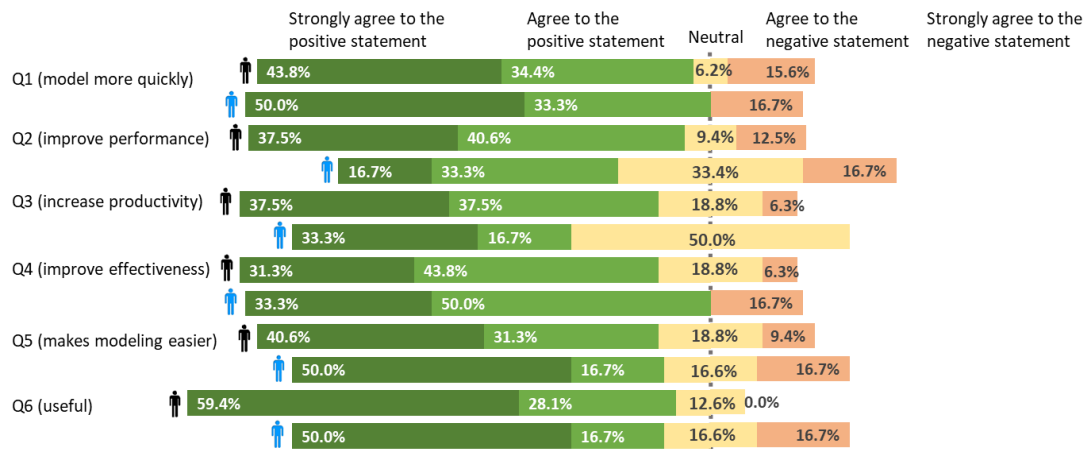


Fig. 13. Perceived usefulness of USINN for interaction and navigation modeling.

The medians of the self-predicted future usage of USINN statements indicated that experienced designers demonstrated a strong intention to use USINN in future when compared to the unexperienced designers. In order to deeply analyze the subjective metrics, we created diverging stacked bar charts illustrating the responses for each statement. Fig. 13 illustrates the perceived usefulness results of the USINN for interaction and navigation modeling.

The results reveal mostly positive perceptions about USINN both for experienced and unexperienced participants and we did not observe a strong agreement to the negative statements when we analyzed Q1 (model more quickly), Q4 (improve effectiveness) and Q5 (makes modeling easier) statements. In relation to Q2 (improve performance) and Q3 (improve productivity) statements, we note that 33.4% and 50% of the experienced designers respectively, indicated neutral responses.

A majority of both experienced and unexperienced participants perceived USINN as a useful notation for modeling interaction and navigation (statement Q6). Some quotes can reinforce this result:

“It is interesting to model interaction and navigation with this model” – P1-C1.

“I can see how the application can answer for specific user actions” – P7-C1.

“(The model) allows us to get a better view of the user-system interaction and to view the user and system, alternative flows, possible interfaces...” – P1-C3.

Regarding the perceived ease of use of USINN for modeling (Fig. 14), we observed the experienced participants provided more positive perceptions about the questions Q1 (easy to learn), Q2 (easy to perform), Q5 (easy to remember) and Q6 (easy to use). On the other hand, we observe a tendency towards neutral responses by unexperienced designers for the same questions.

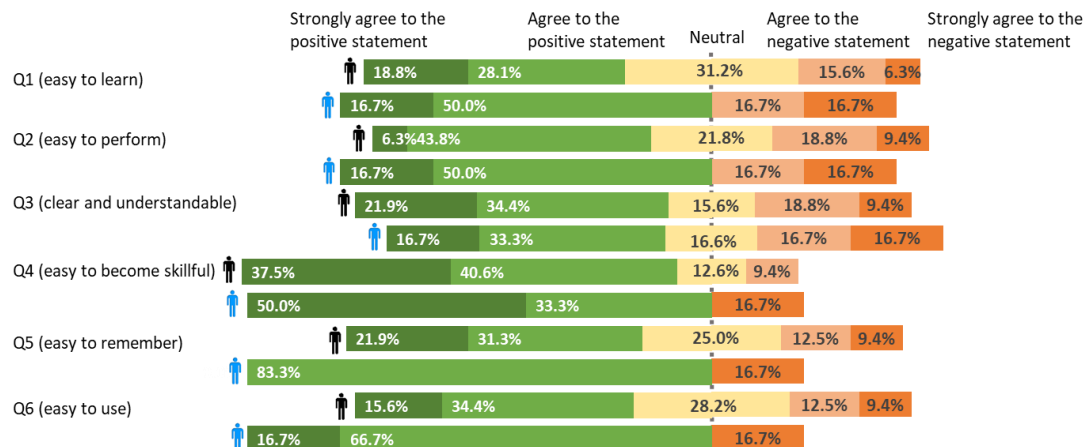


Fig. 14. Perceived ease of use of USINN for interaction and navigation modeling.

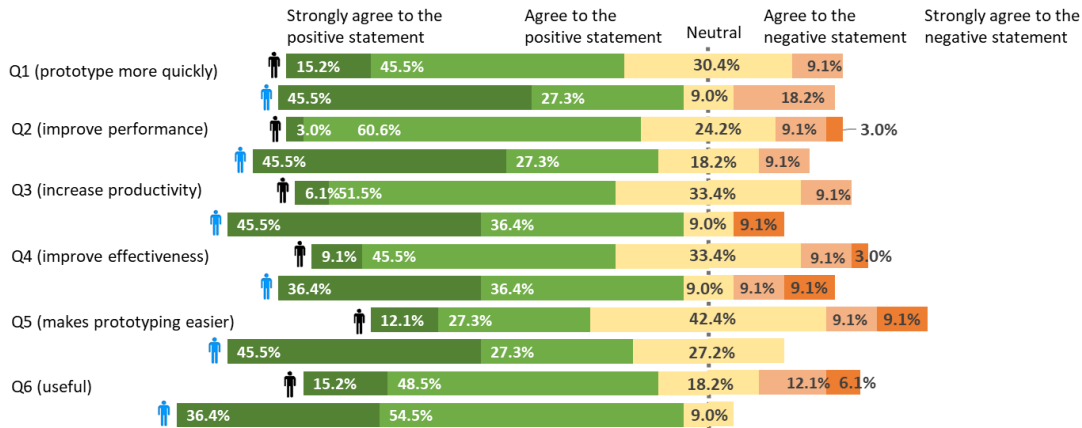


Fig. 15. Perceived usefulness of USINN as basis for creating mockups.

With respect to question Q3 (clear and understandable), we note that experienced designers provided less positive responses than unexperienced designers. When we analyze question Q4 (easy to become skillful), the results for both experienced and unexperienced designers are similar, although the experienced designers did not provide neutral responses. Some quotes from participants indicated features that may have influenced this result:

“It is difficult to represent some aspects related to applications that validate and manipulate the database without updating the interface content” – P7-C1.

“It is easy to learn USINN elements, but it is difficult to model because USINN is complex” – P21-C2.

“The USINN model is easy, because it has few elements in the notation and the elements are clear” – P18-C2.

“I still consider (USINN) a bit confusing, but I think it can be due to the lack of practice” – P29-C3.

“USINN has elements similar to those of UML models. USINN gives us a better view about user interaction” – P42-C3.

“Through the USINN model it is easy to understand the usability requirements, thus leaving the mockup more straightforward and clear” – P27-C2

Fig. 15 illustrates the results of perceived usefulness of the USINN model as the basis for creating mockups. We observed a more positive perception of the experienced participants regarding questions Q2 (improve performance), Q3 (increase productivity), Q4 (improve effectiveness), Q5 (makes prototyping easier) and Q6 (useful for prototyping).

The experienced designers provided fewer neutral responses and more positive responses when compared to the unexperienced designers. The responses of the experienced designers for question Q1 (prototype more quickly) were slightly more positive and with fewer neutral responses when compared to the unexperienced designers.

The following quotes indicate variations in the participants’ perceptions:

“Messages, data collections and flows are the most positive aspects of the USINN model. I had problems to understand specific aspects of the application domain” – P7-C1

With regards to the perceived ease of use of the USINN model as basis for creating mockups (Fig. 16), we noticed the unexperienced participants’ tendency to be neutral regarding the statements. On the other hand, the experienced designers provided more positive responses for all the questions.

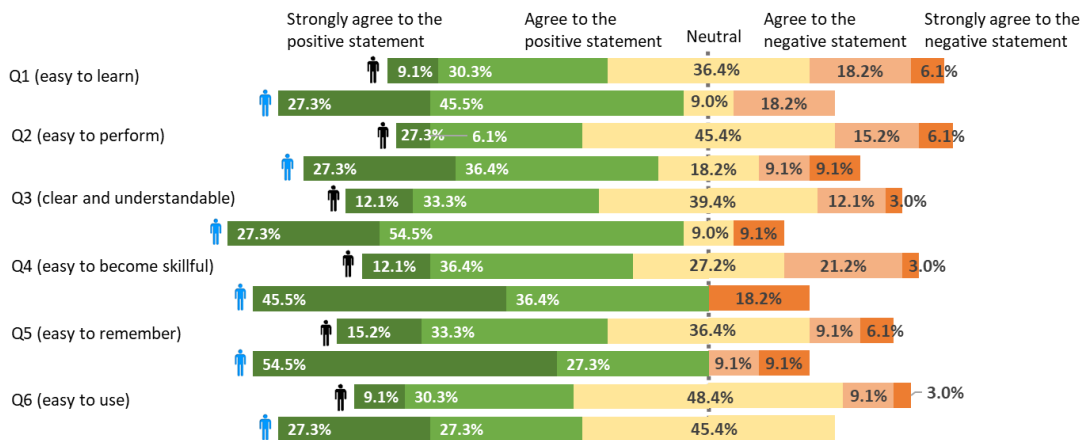


Fig. 16. Perceived ease of use of USINN as basis for creating mockups.

Moreover, the experienced designers provided fewer neutral responses when compared to unexperienced designers, except in relation to question Q6 (easy to use), in which the neutral responses were similar for both experienced and unexperienced designers. Some quotes can explain these results:

“I still have doubts about the USINN diagram, so it was difficult to create the mockups. We also did not know how to represent some requirements” – P3-C1.

“Creating the mockups based only on the USINN diagram is not so effective. I think that when I will create the mockup in a real software industry, I will also use the requirements as basis” – P18-C2.

“As a positive aspect of USINN I can point the ease of understanding the navigation between presentation units and the visual elements to be used in the interface” – P33-C3

“The way in which the model is organized makes me feel lost when following or going back in the interaction” – P22-C2

The self-predicted future usage of USINN for modeling and prototyping is illustrated in Fig. 17. The unexperienced designers provided more positive responses for question Q1 (predict future use) when compared to experienced designers. The majority of the experienced designers did not indicate the intention to use USINN in future. With respect to questions Q2 (prefer to use for modeling) and Q3 (prefer to use for prototyping), the responses of experienced and unexperienced designers were similar. Some quotes are provided by the designers regarding this:

“Unless the practitioner is someone who has practical experience in interaction design and user experience, it is not easy to use USINN in industry” – P3-C1

“I like the USINN model, because I can see how the application can react by certain user actions” – P7-C1

“USINN is a good solution for those who are working in the area and have already the required knowledge” – P15-C2

“I think the USINN model will be essential for future work in my job” – P22-C2.

“USINN is simple and easy to explain for the other professional involved in a project” – P28-C2

“Once you have domain about the USINN elements, it will be easier to use USINN to represent the user interaction” – P37-C3

VIII.DISCUSSION

We previously proposed USINN to increase the usability expressiveness in interactive systems design solutions [8]. USINN enables us to represent the usability mechanisms that existing models partially represent. Our empirical study aimed to investigate whether USINN can support both experienced and unexperienced designers in interaction and interface design.

With our study, we obtained measurements of the quality of the artifacts produced by the participants and the participants’ perception regarding USINN. We obtained correctness values greater than 0.8 for both USINN models and mockups created by the participants. These results indicated that we identified few inconsistency, incorrect fact and ambiguity defects in artifacts. On the other hand, we obtained completeness values under 0.7 in USINN models elaborated by the participants. It indicated that some requirements were not represented in the models. The participants pointed out doubts about the use of the USINN notation and the need for a better knowledge about the notation to effectively apply it.

In order to analyze whether the designers’ experience (having software industry experience or not) affects the results, we decided to compare the quality of the artifacts created by the designers. We observed that the correctness and completeness of models and mockups created by experienced and unexperienced designers were similar.

When we analyze the mean of defects per type of defect and per severity level, we noted that the models created by experienced designers have fewer incorrect fact, ambiguity and omission defects. There was a slight difference between the mean of defects with low and medium severity levels in the models created. The designers represented the requirements in the models, but those models have defects that partially affect the comprehensibility of the requirements.

Concerning the mockups, we did not identify a great difference in the mean of defects per type of defect. Considering the severity level, the experienced designers made more mistakes with medium severity level while the unexperienced designers made more mistakes with high severity level. Thus, the unexperienced designers could have faced difficulties in using the suitable interface components to represent the requirements.

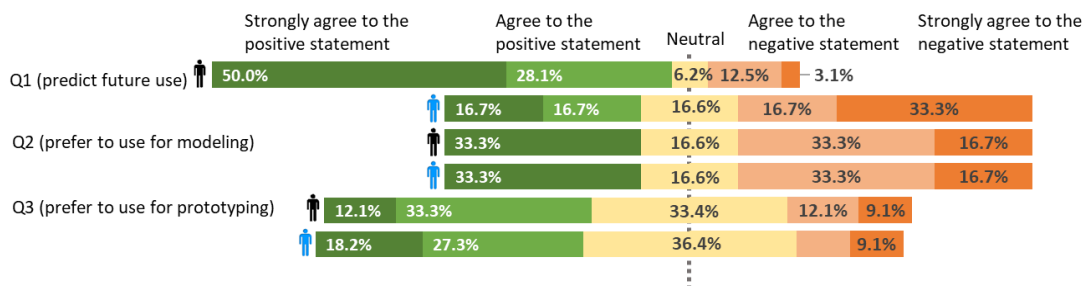


Fig. 17. Self-predicted future usage of USINN.

We collected metrics to evaluate the perceived ease of use, perceived usefulness and self-predicted future usage regarding USINN from the participants' point of view. The experienced designers provided fewer neutral responses when evaluating the perceived usefulness and perceived ease of use of USINN when compared to unexperienced designers. We observed more positive responses from experienced designers with regards to the ease of use for both modeling and prototyping tasks and with regards to usefulness for creating mockups. However, the unexperienced designers provided more positive responses with respect to their future usage intention of USINN. The quotations indicated that the designers considered the experience and knowledge about USINN and usability as factors that can affect the benefits derived from the USINN adoption.

IX. CONCLUDING REMARKS AND FUTURE WORK

In this paper, we presented an empirical study of USINN, a model we proposed to support usability-oriented interaction and navigation modeling. USINN allows representing usability mechanisms underlying the interaction and navigation of interactive systems. Our intention was to evaluate whether USINN can be used in interaction and interface design to create more usable solutions.

To analyze whether the model supports those goals, we conducted an experiment with three undergraduate classes. The quantitative results demonstrate the feasibility of using USINN and the qualitative results indicated the need for improving the ease of use of USINN. Once the experienced designers did not provide much intention to adopt USINN, we intend to carry a case study in industry to investigate how USINN can be adapted to a software development lifecycle.

To deeply investigate the difficulties the participants pointed out, we will conduct an observational study and evaluate USINN based on the Cognitive Dimensions Notation (CDN) as discussed in [23]. Additionally, we aim to conduct an empirical study with design professionals who are involved in systems development and interaction design in industry.

Also, we intend to propose (i) guidelines to support the inclusion of usability mechanisms when novice designers or non-usability specialists adopt USINN; (ii) design patterns for interaction and navigation modeling with USINN focusing on usability mechanisms.

ACKNOWLEDGMENTS

The authors acknowledge the financial support granted for this research by CAPES process 175956/2013 and CNPq processes: 309828/2015-5 e 423149/2016-4.

REFERENCES

- [1] Adriana Lopes, Anna B. Marques, Simone D. J. Barbosa, S, and Tayana Conte. 2015. Evaluating HCI Design with Interaction Modeling and Mockups: A Case Study. In Proc. of International Conf. on Enterprise Information Systems (ICEIS 2015), 79-87. DOI= 10.5220/0005374200790087.
- [2] Andy Gimblett and Harold Thimbleby. 2010. User Interface Model Discovery: Towards a Generic Approach. In Proc. of Engineering Interactive Comp. Systems (EICS'10), 145-154. DOI= 10.1145/1822018.1822041.
- [3] Ugo B. Sangiorgi, and Simone D.J. Barbosa. 2009. MoLIC Designer: Towards Computational Support to HCI Design with MoLIC. In Proc. of Engineering of Interactive Comp. Systems (EICS'09), 303-307. DOI= 10.1145/1570433.1570489.
- [4] Michel Beaudouin-Lafon. 2004. Designing interaction, not interfaces. In Proc. of Working Conf. on Advanced visual interfaces (AVI'04), 15-22. DOI= 10.1145/989863.989865.
- [5] Natalia Juristo, Ana M. Moreno, and María-Isabel Sanchez-Segura. 2007. Analysing the impact of usability on software design. J. Syst. Software 80, 9: 1506-1516.
- [6] Laura Carvajal, Ana M. Moreno, María-Isabel Sanchez-Segura, and Ahmed Seffah. 2013. Usability through Software Design. IEEE T. Software Eng 39, 11: 1582-1596.
- [7] Jose I. Panach, Natalia Juristo, Francisco Valverde, and Oscar Pastor. 2015. A framework to identify primitives that represent usability within Model-Driven Development methods. Inf Softw Technol 58: 338-354.
- [8] Anna B. Marques, Simone D. J. Barbosa, Tayana Conte. 2017. Defining a notation for usability-oriented interaction and navigation modeling for interactive systems. SBC Journal on Interactive Systems 8, 2: 35-49..
- [9] Werner Gaulke and Jürgen Ziegler. 2016. Rule-Enhanced Task Models for Increased Expressiveness and Compactness. In Proc. of Engineering Interactive Comp. Systems (EICS'16), 4-15. DOI= 10.1145/2933242.2933246.
- [10] Giorgio Brajnik and Simon Harper. 2016. Measuring Interaction Design before Building the System: a Model-Based Approach. In Proc. of Engineering of Interactive Computing Systems (EICS'16), 183-193. DOI= 10.1145/2933242.2933246.
- [11] Simone D. J. Barbosa, and Maíra G. de Paula. 2003. Designing and Evaluating Interaction as Conversation: a Modeling Language based on Semiotic Engineering. In Proc. of Int. Workshop on Design, Specification, and Verification of Interactive Systems (DSV-IS'2003). 16-33. DOI= 10.1007/978-3-540-39929-2_2.
- [12] Fernando Molina, and Ambrosio Toval. 2009. Integrating usability requirements that can be evaluated in design time into Model Driven Engineering of Web Information Systems. Adv Eng Softw 40, 12: 1306-1317.
- [13] Eric Barboni, Jean-François Ladry, David Navarre, Philippe Palanque and Marco Winckler. 2010. Beyond Modelling: An Integrated Environment Supporting Co-Execution of Tasks and Systems Models. In Proc. of Engineering of Interactive Comp. Systems (EICS'10), 165-174. DOI= 10.1145/1822018.1822043.
- [14] Magalí González, Luca Cernuzzi, Oscar Pastor. 2016. A navigational role-centric model oriented web approach MoWebA. Int. J. Web Eng. Technol 11, 1: 29-67.
- [15] Natalia Juristo, Ana M. Moreno, and María-Isabel Sanchez-Segura. 2007. Guidelines for Eliciting Usability Functionalities. IEEE T. Software Eng 33, 11: 744-758.
- [16] Anna B. Marques, Simone D.J. Barbosa, Tayana Conte. 2016. Artifacts used in empirical studies in the research USINN: A Usability-oriented Interaction and Navigation model, 27-32. Available on <https://goo.gl/bQwA9r> (In Portuguese).
- [17] Maria F. Granda, Nelly Condori-Fernández, Tanja E. J. Vos, and Oscar Pastor. 2015. What do we know about the defect types detected in conceptual models? In Proc. of Int. Conf. on Research Challenges in Information Science (RCIS 2015), 88-99, DOI= 10.1109/RCIS.2015.7128867.
- [18] Viswanath Venkatesh, and Hillol Bala. 2008. Technology Acceptance Model 3 and a Research Agenda on Interventions. Decision Sciences 39, 2: 273-315.
- [19] Leonard Marascuilo and Ronald Serlin. 1988. Statistical Methods for the Social and Behavioral Sciences. Journal of Educational Statistics 15, 1: 72-77.
- [20] Sidney Siegel and N. John Castellan. 1988 Nonparametric Statistics for the Behavioral Sciences (2nd ed.) McGraw-Hill International Editions.
- [21] Edward Carmines and Richard Zeller. 1979. Reliability and Validity Assessment. SAGE Publications.

- [22] William King and Jun He. 2006. A meta-analysis of the technology acceptance model. *Inf Process Manag* 43, 6: 740-755.
- [23] Juliana J. Ferreira, Clarisse S. de Souza, Renato Cerqueira. 2014. Characterizing the Tool-Notation-People Triplet in Software Modeling Tasks. In *Proc. of Brazilian Symp. on Human Factors in Comp. Systems (HC'14)*, 31-40.