Assessing the Semiotic Inspection Method – The Evaluators' Perspective

Soraia de Souza Reis [1]
Raquel Oliveira Prates [1]

Abstract: This paper presents an assessment of the Semiotic Inspection Method aimed at understanding its costs, benefits, advantages and disadvantages from the evaluators' perspective. We applied a questionnaire to novice evaluators and interviewed the authors of the method (representing the experts' evaluators). An analysis of the responses shows interesting insights and characteristics of the method.

1 Introduction

During the design process of an interactive system it is important to evaluate its interface against relevant quality properties [1]. Systems' interface assessments are typically performed using evaluation methods which aim at supporting evaluators in assessing interface properties and identifying interaction problems [1]. Nowadays, there are a large number of evaluation methods that can be used to assess interfaces and they differ in several aspects. Some of the most consolidated and well known methods within the HCI community are empirical methods, which are based on the best practices defined by experts over many years of study, such as: Heuristic Evaluation and Usability Testing. The disadvantage associated with empirical methods is that the interaction with computer systems is constantly changing. Therefore, it becomes increasingly difficult to use an applied knowledge for the development of interfaces, since the problem to which it should be applied to is intrinsically unstable [2]. Moreover researchers have argued that these methods may not always be the most appropriate choice to assess the contribution of a system or research and new methods should be investigated [3].

For this reason, theoretical approaches in HCI have been gaining power and the need for investing and advancing research in the area has been recognized as necessary in order to contribute to the quality improvement of information technology artifacts [4]. The advantage associated with the use of theoretical methods is that they involve epistemological, ontological and methodological aspects that allow us to analyze the object of study [2]. In other words, they allow computer interface designers to explain the effects that certain design choices can cause during user-system interaction. The Semiotic Inspection Method (SIM) (our study object), is one of these methods.

[1]Departamento de Cincia da Computao – Universidade Federal de Minas Gerais
{soraiareis,rprates@dcc.ufmg.br}

SIM was proposed in 2006 [5] and, although SIM has been used in different domains since its proposal [6], only recently researchers have invested in demonstrating its benefits. Some studies have investigated its applicability to specific domains (e.g., educational domain) [7, 8, 9]. More recently, a broader assessment of its applicability was conducted [6] and has shown a great advantage of the method: its independency of technology and domain. These results make it even more relevant to continue evaluating it.

SIM is still a recent theory-based method and in order to better understand its costs and benefits, as well as those of theory-based methods in general, it is important to understand how evaluators perceive and use it [10]. An assessment from the evaluators' perspective is important because evaluators are responsible for deciding which method to use for an interactive system evaluation and identify the costs and benefits from the evaluators' perspective may be important in defining (or revising) a strategy to present and/or teach the method.

Therefore, in this paper we present an assessment of SIM from the evaluators' perspective aiming at identifying their perception about SIM's costs, benefits, advantages and disadvantages. In order to do so, we applied a questionnaire to novice evaluators and made interviews with the authors of SIM (which represent the expert evaluators) in order to compare, contrast and consolidate their views. Although we have not obtained a definite and final set of indicators, interesting insights were identified.

This paper is organized as follows. Next we present the Semiotic Engineering Theory and the Semiotic Inspection Method. Then we present the related works followed by the assessments conducted and its results and analysis. In Section Discussion we present a consolidated view of the advantages, disadvantages, costs and benefits of SIM from the evaluators' perspective. Finally we present the final considerations and future works.

2 Theoretical Framework

2.1 Semiotic Engineering

Semiotic Engineering (SemEng) [2] is an explanatory theory of HCI that perceives the system's interface as a message being sent from designers-to-users. This message conveys to users the designers' understanding of whom they are, their goals, and how they can interact with the system to achieve these goals. The interface is considered a meta-communication artifact, since the designer-to-user communication takes place through user-system interaction.

In a semiotic perspective a message is composed of signs, which can be defined as anything that represents something for someone [11]. In SemEng the sign can be classified into three classes [12, 5, 13]: metalinguistic, static and dynamic. The metalinguistic signs are signs that designers use to explicitly communicate to users the meanings encoded in the

system and how they can be used (e.g., tips, help). Static signs are signs whose meaning is interpreted independently of temporal and causal relations, in other words, the interpretation is limited to the elements that are on the interface at a single moment in time (e.g., button's state). The dynamic signs represent the system behavior, in other words, they can only be perceived through the interaction with the system (e.g., action triggered by a button).

Within the SemEng theoretical framework, a relevant interface property is its communicability which is defined as a distinctive quality of interactive systems that communicate efficiently (in an organized and resourceful way) and effectively (achieving the desired result) to users its underlying design intent and interactive principles [14]. Based on the SemEng theory, methods to evaluate the interfaces communicability have been proposed: Semiotic Inspection Method (SIM) [12, 5, 13], Communicability Evaluation Method (CEM) [2, 12, 14] and the Intermediated Semiotic Inspection Method (ISIM) [9]. In this paper our focus is in assessing SIM, which we next present in detail.

2.2 Semiotic Inspection Method

SIM is an inspection method in which evaluators assess the communicability of interactive computer-based artifacts. The goal is to examine the meta-message being sent by the designer (i.e., system's interface) to users, identifying potential breakdowns users may experience and reconstructing the intended meta-message [12, 5].

Before SIM's application, evaluators must define the system's scope to be examined and the user's profiles being considered in the evaluation. The scope and profiles should be used to generate evaluation scenarios that will provide the contextual structure necessary for the analysis.

SIM is comprised of five steps [12, 5, 13]: (1) analysis of metalinguistic signs through system's documentation and help; (2) analysis of static signs; (3) analysis of dynamics signs. In each of these steps, the meta-message is reconstructed based solely on the signs analyzed, and potential breakdowns are identified. The next two steps then follow: (4) comparison and contrast of the meta-message identified in steps (1), (2) and (3); and finally (5) reconstruction of a unified meta-message and final assessment of the system's communicability. Steps (1), (2) and (3) are done iteratively. In these steps the evaluator makes a segmented analysis of the system, one for each of the three classes of signs. This segmented analysis allows the evaluator to inspect in detail what and how the designer communicates with each of these types of signs. In steps (4) and (5), the evaluator reconstructs the final meta-communication message by comparing, integrating and interpreting the data collected in previous steps of the method. At this point, the evaluator is able to articulate her findings about the system's communicability.

The method can be applied in two different contexts: technical and scientific [5, 13].

In technical contexts, it can be used to improve the quality of designer-user communication. The aim, focus, and other circumstances of the evaluation are dictated by business needs, as well as industrial or commercial interests. In scientific contexts, in the other hand, the purpose, focus, and the circumstances of evaluation are dictated by the research questions and methodological consistency. The immediate goal of using it in the scientific context is primarily to advance knowledge.

3 Related Work

The introduction of new methods with different approaches and the lack of understanding of their capabilities and limitations have intensified the need to determine which methods to use in certain situations [15]. In order to provide useful guidance in instructing evaluators on costs and benefits regarding a method, studies should be conducted outlining its advantages and disadvantages [16].

In the usability context, once methods had been proposed studies aiming at assessing those methods emerged. Taking as an example the Heuristic Evaluation (HE) and Usability Testing methods (the most consolidated usability methods) assessments emerged with different purposes. HE was initially assessed in relation to the severity and number of problems encountered, in terms of the number of evaluators needed to find most of the problems, and the influence of evaluators' experience [17, 18]. Then, comparative studies were performed and presented as the work done by [19] (comparing the methods: HE, guidelines, Cognitive Walkthrough and Usability Testing); [20, 21] (comparing the impact of different evaluator profiles in the application of HE and Cognitive Walkthrough); and [16] (compared Usability Testing with individual and team walkthrough methods). More recently, comparison of methods continue to be conducted, however to investigate its application (or a proposed adaptation) to specific domains. For instance, [22] presents a comparison of a user-based method with an inspection-based method using a set of groupware heuristics to evaluate groupware systems and [23] presents a web tool evaluation for nurse scheduling using HE and Think-Aloud.

Despite the efforts to provide information about interface evaluation methods, some researchers criticized the way in which some of the usability methods assessments were conducted [24]. In this context, motivated by the lack of criteria to assess methods, some researchers [15] proposed measures to assess usability evaluation methods. The proposals present measures for assessing the performance of the methods in a comparative way such as validity, thoroughness and reliability. However, [25] argues that, although these measures are mandatory for researchers seeking to establish the efficacy of a given procedure, the measures thoroughness and efficiency are meaningless and impossible to calculate, thus, irrelevant in HCI practice. The procedures and measures used recently are still being criticized [10], and,

therefore, this indicates that the HCI field still needs further studies related to methodologies to support researchers to conduct assessments of the methods that have emerged recently, as well as those that take into account properties other than usability.

SemEng methods have been proposed recently and criteria applied to comparing usability evaluation methods may not be the most appropriate for communicability evaluation methods [26]. Since the methods are recent, only a few studies have been conducted to evaluate them. A study discussing the cost-benefits of CEM in comparison to well-known usability evaluation methods discusses relevant criteria to compare these methods and presents the results [26]. The ISIM was proposed and evaluated in regard to its contributions to educational systems evaluations [9].

Regarding SIM, some studies have investigated its applicability to specific domains, namely educational domain [9], collaborative systems [8] and Human-Robot Interaction [7]. All these works have shown that the method could be successfully applied in the specific domain being investigated. Recently, a broader assessment of the applicability of SIM was conducted indicating that SIM has been applied without adaptations to different technologies and domains, and in each situation it has been able to successfully identify specific issues [6]. Thus, authors argue that this can be perceived as an advantage of SIM over other methods that usually need to be adapted in order to take into consideration specificities of the domain or technology. This first result makes it even more relevant to continue evaluating the method.

Thus, in this paper we performed an assessment aiming at identifying costs and benefits of SIM from the evaluators' perspective. To do so, we applied a questionnaire to novices' evaluators and interviewed expert evaluators who had used SIM. The motivation for evaluating SIM from the evaluators' perspective is that evaluators are the users of the method, and the adoption of the method involves people being able to apply it and realize its value. Next we present how this assessment was carried out and the results obtained.

4 The Assessment

The assessment from the evaluator's perspective was conducted in two steps. In the first step we applied a questionnaire to students and researchers who had used SIM at least once with the goal of obtaining initial information on the characteristics of the method from novice evaluators' perspective. In the second step we conducted interviews with the authors of the method. The reason for having done the interviews with the authors was due to the fact that they are the most experienced people with the method and, therefore, represent the expert evaluators' perspective. The overall goal of this assessment was to compare, contrast and consolidate their (i.e., novices and experts) perception about the costs, benefits, advantages and disadvantages of SIM.

4.1 The Survey

The survey was conducted through a questionnaire containing **31** questions divided between multiple-choice and open-ended questions. Multiple-choice questions were mandatory and open-ended questions were optional. The questions were divided according to their focus as follows:

- **Participants' profile:** 10 questions regarding general information (e.g., gender or age), training and professional experience;
- **Learning:** 2 questions to identify the courses in which participants had learned SIM;
- **Experience with SIM:** 15 questions about participant's experience in applying SIM, and the challenges experienced in learning and applying the method; and
- **Experience with HCI:** 4 questions about participants' experience using other HCI evaluation methods.

The reason for using a questionnaire was that it would be possible to reach a larger number of people, including evaluators who authors of this study did not know of. The questions were designed to allow us to collect information about SIM, as well as to have an overview of the general evaluation experience of the participants. Regarding SIM the idea was to see which steps were considered difficult (if any) in the application and analysis of the method, and also collect the participants' opinions regarding advantages and disadvantages and any other aspects they considered relevant about the method (through open-ended questions).

The questionnaire was applied from June 29^{th}, 2011 to July 29^{th}, 2011 and distributed by e-mail to: (1) the Brazil national HCI e-mail list, since to the best of our knowledge, at the time of the research, all publications regarding SIM and its use involved at least one Brazilian researcher[2]; and (2) HCI researchers and professors who were known to have worked with the Semiotic Engineering theory asking them to distribute the questionnaire to their students and other people who they knew had applied SIM. It is noteworthy that the authors of the method were asked not to answer the questionnaire. The questionnaire was answered by 25 participants. We considered a good number of responses since SIM is a relatively recent method, and it is not widely used or taught (yet). Besides, although the number of responses is small, we believe that this initial research is necessary to raise the main issues regarding the method to be further investigated using deeper approaches (e.g., empirical assessments).

[2] At the time this research was conducted all the studies with the method had the participation of at least one Brazilian. However, when updating the systematic literature review ([6] – described in the previous section) a publication by a group not from Brazil in which the method was applied was identified [27], which can be taken as an evidence of a broader adoption of the method.

4.1.1 Participants Profile Among participants, **12** were female and **13** male. Their age varied from 19 to 40, but most (**16**) were between 19 and 25 years old. Professionally[3], more than half of the respondents were students (**16**), **11** were professionals in the Information Technology (IT) field; **4** worked as researchers; and **1** was a professor at a university. Only **8** participants have professional experience (i.e., working in industry as professionals or interns) in the HCI field. Regarding their highest educational level **15** were undergraduate students or have a university degree; **8** had a MSc degree or were a MSc candidate; and **2** were PhD candidates.

Participants mostly (**16**) had their current education level in Computer Science, **8** in Information Systems and only **1** person said that was studying both Information Systems and Computer Science. Participants were from different places, where most of them – **15** – studied or were attending a course at the *Universidade Federal de Minas Gerais*. The other participants are from Brazilian universities in other states, such as *Paran* and *Rio Grande do Norte*, or abroad.

Graduate students (MSc and PhD) were asked to inform their research area[4]: **8** participants did research in HCI field and **2** in Software Engineering. Although not requested, some participants also mentioned their research sub-area, which were Learning Support Systems, Learning Objects and Games.

Among participants, **5** used SIM once; **16** used SIM twice; **2** used SIM three times; and **2** used SIM four or more times. All participants had also applied other evaluation methods: **23** applied Communicability Evaluation Method (another Semiotic Engineering evaluation method); **22** applied Heuristic Evaluation; **18** applied Usability Test; **5** applied Cognitive Walkthrough; and **4** applied Think Aloud Protocol. In this question participants could not mention other methods. It is noteworthy that most participants (**24**) had experience with these methods using them in a course context; **12** used for research purposes; and **7** applied them professionally.

4.1.2 Results and Analysis The analysis of the questionnaire (as well as the analysis of the interview) aimed at identifying the evaluators' perception about SIM's costs, benefits, advantages and disadvantages. The analysis performed is of a qualitative nature. The reason for conducting a qualitative analysis is that the goal of the research was not to get to a conclusion regarding the cost-effectiveness of the method, but rather understand, compare, contrast and consolidate the main difficulties and issues being experienced by evaluators applying the method (i.e., raise their perception on the main costs and benefits of the method).

[3]Participants could choose more than one option.
[4]Participants mentioned more than one research area.

Costs and Effectiveness Analysis The questionnaire participants assessed the difficulty in applying SIM on a scale from 1 to 5 (with 1 being very easy and 5 very difficult). Most participants evaluated the method as Medium or Difficult to apply. Participants were also asked about which step they believed were the most difficult (if any). Among the steps of the method, the fourth step was the one that participants found the hardest (**48%**) . The fourth step is the one in which the evaluator compares the meta-messages obtained by the analysis of each sign class (steps 1 to 3), consolidating them in a single meta-message and identify inconsistencies among them. If evaluators are not able to perform well the fourth step, they may not be able to identify relevant problems that could result from contrasting the meta-message generated from the analysis of each sign class.

They were also asked what was their greatest difficulty in applying the method (if any): difficulty to learn the method; difficulty to understand the theory that underlies the method; or difficulty related to time and concentration required to conduct a good evaluation. They also had the option to answer that they had no difficulty or mention another difficulty they might have experienced. All available choices were selected by at least 1 participant. However, two of them stood out significantly: (1) time and concentration required; and (2) learning the theory. Almost half the participants indicated that their major difficulty was related to the time and concentration required. This is an indication that for these evaluators the time and concentration needed in the evaluation is a cost to be considered. However, by looking at the profile of the evaluators who chose this option, we notice that the more experienced SIM evaluators (i.e., applied SIM more than 4 times), which were two participants, did not point this option as a problem. They mentioned to have no difficulty with SIM. Thus, this could be an evidence that as evaluators become more experienced with SIM, they either take less time to apply it or make a better use of the time spent.

About one third of the participants indicated that the underlying theory is their biggest challenge in applying the method. Theory based methods usually require an understanding of the underlying theory, so this could be expected. The good news is that once evaluators learn the theory they would not face this challenge any longer. Furthermore, learning the theory could be motivated by reasons other than just applying SIM. As mentioned, participants who applied SIM more than 4 times mentioned not having difficulties with the method, and this reinforces our argument that the difficulties decreases as the evaluators become more experienced.

Due to the requirement that understanding the theory is necessary to apply the method, participants were asked how important they felt the knowledge of Semiotic Engineering theory was in the ability to apply SIM (on a scale from 1 to 3, with 1 being low and 3 high). All participants believed the theory was necessary or useful in the application of the method. The great majority (over **70%**) believed that the theory was essential in achieving good results, whereas the others thought it was not essential, but could help. In analyzing this result

combined with the previous question, we notice that all of them understand the importance of knowing Semiotic Engineering has on the method, but do not consider learning the theory the major cost of applying SIM.

Regarding the effectiveness of using the method, participants were asked how they evaluated the relevance of the problems identified using the method. They had three options and could select one or more of them: SIM helped in identifying problems that had no impact in the system's use; SIM helped to formalize problems that had been identified before; and SIM helped in identifying relevant problems. Almost all participants (**92%**) answered that SIM helped finding relevant problems. In addition, **72%** believed that SIM also allowed the formalization of known problems. None of them reported that the method could only find problems that did not impact the system's use.

In conclusion, we can mention that in the evaluators' perspective the costs of SIM are mainly related to (1) high time/effort demanded in applying the method; and (2) knowledge and skills needed to do so, since knowledge of Semiotic Engineering theory is required. Regarding the effectiveness of SIM, we can conclude that it is an effective method since it allows identifying relevant problems of the system being evaluated and formalizing perceived problems.

Advantages and Disadvantages Analysis We now present the advantages and disadvantages of SIM from the novices' evaluators' perspective. This information was identified based on five optional open-ended questions. The participants were asked to:

- comment whether they would or not apply SIM in other interface evaluations (22 responses);
- list the main advantages of SIM (23 responses);
- list the main disadvantages of SIM (21 responses);
- list SIM's costs and benefits related to other methods (20 responses); and
- make any other comments (5 responses).

We used Grounded Theory [28] techniques to analyze the responses. We present in Table 1 and Table 2, respectively, a summary of the advantages and disadvantages encountered. We also present in the column "#" the number of participants that mentioned it. As mentioned earlier, the analysis is of a qualitative nature and the reason for presenting the number of participants who reported an advantage/disadvantage is only to illustrate which issues had more or less agreement upon.

Table 1. SIM advantages.

	Advantages	Participants	#
A1	Identifies problems related to communicability.	P5; P15; P16; P22; P24; P25	6
A2	Allows the analysis of the impact of each sign class (i.e., metalinguistic, static and dynamic) independently.	P2; P9; P12; P18; P20	5
A3	Allows formalizing, arguing and explaining the problems encountered while interacting with the system.	P1; P4; P20	3
A4	It has a good cost-benefit relation (i.e., good results at a relatively low cost).	P4; P6; P17	3
A5	Only one evaluator is needed to perform an evaluation.	P17; P22; P24	3
A6	It is a theory-based method.	P5; P18; P25	3
A7	It is a simple method to learn/understand and apply.	P6; P11; P17	3
A8	It is a method that produces good results and allows finding important problems.	P3; P7	2
A9	Allows a thorough analysis of the system and identification of its problems.	P1	1
A10	Allows generating new knowledge using a scientific application of SIM.	P7	1
A11	It does not require an expert evaluator to understand the problems reported.	P1	1
A12	It is a method that, unlike other interface evaluation methods, performs an assessment of the help and system's documentation.	P2	1
A13	Allows identifying the impact of problems related to other properties of the system (i.e., accessibility, sociability).	P5	1
A14	Can be used to evaluate system's interface in the beginning of the development process.	P10	1
A15	Allows evaluating the system from evaluators' point of view (i.e., meta-message emission).	P25	1
A16	It can be applied to different domains and technologies without adaptations.	P5	1
A17	It is a low cost method (i.e., does not demand too much time and effort).	P22	1
A18	It is a good guide to find problems (i.e., provide procedures that guide the evaluator during the evaluation).	P17	1

Analyzing Table 1 we can see that according to participants, the major advantage of SIM is that it allows identifying problems related to communicability **[A1]**. Although communicability evaluation is SIM's main goal, it shows that participants understand it as a relevant property of an interactive system. Also in **[A2]** participants list as an advantage the fact that the method can be used to analyze the impact of each sign class independently. This advantage may be directly related to **[A9]** in which a participant states that SIM allows a thorough analysis of the system, even though participants themselves did not indicate this relationship.

Another advantage of SIM is associated with the fact that it allows formalizing, arguing and explaining the problems encountered **[A3]**. We understand that any evaluation method allows evaluators to explain the problems encountered. However, one participant (P20) explicitly stated that:

> *"To formalize analysis (getting more concrete arguments against the problems under consideration); as well as helping to define more precisely where the interaction and communicability problem of the system is (it is possible to define if a problem occurred because of a particular static sign and not by others, for example)."*[5]

This statement shows the advantages related to how SIM leads evaluators to describe the problems due to its steps and to the underlying theory. Thus, **[A3]** could be closely related to the advantage **[A6]** in which participants stated that an advantage is being a theory-based method, even though they did not explain why they believed this was an advantage. The fact that participants have perceived through the application of the method that it has benefits due to its theoretical basis indicates that participants' views strengthen the researchers' claims that theoretical approaches in HCI are needed [4, 2]. In addition, the advantage **[A6]** may also be related to **[A10]** where participants think that the possibility of generating new knowledge using a scientific application of SIM is an advantage.

The participants mentioned that SIM does not require an expert evaluator to understand the problems reported **[A11]**. This advantage could be argued to apply to any method, however, since SIM requires the learning of Semiotic Engineering theory in order to apply the method, the fact that it is possible to generate a report that does not necessarily require this knowledge to be understood by the reader could be perceived as an advantage.

At last, the participants mentioned the applicability of SIM to different domains as an advantage of the method **[A16]**. Their view is aligned with the investigation presented in the

[5]The original statement (translated by the authors) in Portuguese by P20 was: *"Formalizar anlises (obtendo argumentos mais concretos contra os problemas analisados); alm de ajudar a definir com mais exatido onde o problema de interao e comunicabilidade do sistema (possvel definir que um problema ocorre por causa de um determinado signo esttico e no por outros, por exemplo)"*.

previous section that showed that SIM can be applied to different domains and technologies without adaptations [6].

The biggest drawback raised by participants (shown in Table 2) is that SIM is a method that demands a lot of time and effort to be applied **[D1]**. This disadvantage is related to **[D6]** where participants stated that SIM requires the generation of an extensive and detailed documentation which increases the evaluation time, and also to **[D8]** where participants stated that the method is laborious, repetitive and tiring at times which may cause the evaluator to overlook some problems. These three disadvantages may be some of the factors that explain one of the main costs identified in the closed questions: time and concentration required by the method. Also, they could justify the need to offer an evaluation supporting tool. The inexistence of such tool has also been identified as a disadvantage **[D5]**.

Table 2. SIM disadvantages.

	Disadvantages	**Participants**	**#**
D1	Demands a lot of time and effort to apply it.	P5; P7; P10; P12; P20; P23	6
D2	It is based on a complex theory.	P2; P5; P9; P16; P17; P18	6
D3	It is a difficult method to apply with a high learning curve.	P2; P4; P14; P22; P25	5
D4	The experience of the evaluator is very important to get good results. Novices may have difficulties and may not generate good results.	P3; P17; P23	3
D5	There is no tool available to support the evaluation.	P1; P5; P18	3
D6	It is a method which requires the generation of an extensive and detailed documentation, which increases the evaluation time.	P1; P7	2
D7	Lack of material to support learning, such as case-study examples on how to apply the method.	P2; P18	2
D8	It is a laborious and repetitive method, and is tiring at times. This may cause the evaluator to overlook some problems.	P11; P20	2
D9	It has a low cost-benefit relation.	P7	1
D10	Lack of experts, which makes difficult to arrange a team to apply the method.	P12	1

In addition, participants said that SIM is a method based on a complex theory **[D2]**. Although participants did not make associations between disadvantages, we can notice that this could probably explain the disadvantage that states that it is a difficult method with a high

learning curve **[D3]**. A participant's (P9) statement *"Difficult, too theoretical."*[6] explicitly creates the link between the two. Again these disadvantages reinforce the cost regarding knowledge and skills necessary to apply the method that was mentioned previously.

Regarding learning the method, most of the participants said to have learned it through classes and educational materials. However, in **[D7]** participants mentioned the lack of material to support learning the method as a disadvantage. This probably indicates that the existing materials are not sufficiently thorough or complete. The deficiency of the existing material could also increase the cost of learning the method, pointed by participants.

Another disadvantage mentioned by the participants is related to the influence of evaluators' expertise in the results **[D4]**. Although no study of the kind has yet been performed for SIM, we would argue that expertise could influence positively the results, especially taking into account that the knowledge of the theory has already been identified as necessary in applying the method and getting good results. Furthermore, a participant cited that there is a lack of experts in the method, which makes it difficult to arrange a team to apply the method **[D10]**. The participant did not offer any comments regarding the context in which he experienced these difficulties. At any rate, SIM is a recent method and it requires time before any method could have a large number of experts. This disadvantage also justifies the need of studies to assess the method such as the ones being presented in this work which allows people to better understand its costs and benefits before investing in learning.

By contrasting the results found, we notice that participants do not always agree on the advantages and disadvantages of the method. This is the case, of the advantage stated by some participants that the method is easy to learn **[A7]** and the disadvantage that it has a high learning curve **[D3]**, as well as the cost of the required knowledge to apply the method. By looking at the answers of those participants who mentioned **[A7]**, they were consistent along the questionnaire in answering that for them SIM is not difficult. This could be a positive result, since it may indicate that for some people the method is easy. However, a deeper investigation is needed to identify if the causes of these differences are personal learning-style, or due to their experience with the method, for instance evaluating a simple versus complex system.

The second contradiction was regarding the cost of the method (among **[A17]**, **[D1]** and the analysis of the closed questions presented previously). Only one participant (P22) mentioned that SIM does not demand too much time and effort:

> *"It is a very good method to analyze the communicability of a system and it requires less time and resources compared to others."*[7]

[6] Original statement in Portuguese by P9: *"Difcil, muito terico."*

[7] Original statement in Portuguese by P22: *" um mtodo muito bom na anlise da comunicabilidade de um sistema e que demanda pouco tempo e recursos comparados a outros mais."*

The answers given by the participant were consistent with this view along the questionnaire. In the question in which participants were asked what was their greatest difficulty in applying the method he answered that he had difficulty to understand the theory that underlies the method and did not check the option about the difficulty related to time and concentration required. Notice that in his/her answer he/she mentioned that SIM requires less time compared to other methods. However, he did not mention which methods he was comparing SIM to, but indicated in this questionnaire that he had experience with HE and CEM. This contradiction shows that further investigation of the cost associated with the time/effort is needed, perhaps involving the measurement of the time spent by different users' profiles.

Finally, the third contradiction occurred between **[A4]** and **[D9]** regarding SIM's cost-benefits. Only one participant (P7) mentioned that SIM has a low cost-benefit ratio **[D9]** and he stated the following:

> *"I believe that the communicability evaluation of a system could be made somewhat more superficial, placing greater emphasis on other qualities of use such as usability. After all, a system may have high communicability and still have low usability. Therefore, by being a method that evaluates only the communication, in particular, only the communication emission, I believe that the effort to apply it is not justified."*[8]

We can argue that communicability is a property that was proposed to add to the HCI field and never intended to substitute usability [2]. In other words, they are two distinct properties that ideally should be combined to achieve a better quality of use in interactive systems. Therefore, the fact that it does not evaluate properties that it was not meant to evaluate should probably not be considered a cost of the method. We understand that in real contexts evaluators may not be able to conduct more than one evaluation due to available resources, and will have to take into account whether communicability or usability is more relevant to the system being evaluated. However, deciding the goal of the evaluation and the specific questions to be answered are steps to be taken before deciding the method to apply, as has been described in the DECIDE framework [1]. Furthermore, although we could infer that participant P7 may not think that the communicability property is as relevant as usability; other participants have mentioned being able to evaluate communicability issues as an advantage **[A1]** and that it has a good cost-benefit relation **[A4]**. In addition, participants have also mentioned that SIM produces good results and allows the identification of relevant problems **[A7]**.

[8]Original statement in Portuguese by P7: *"Acredito que a avaliao da Comunicabilidade de um sistema poderia ser realizada de forma um pouco mais superficial, dando maior nfase a outras qualidades de uso, como a Usabilidade. Afinal, um sistema pode ter alta Comunicabilidade e ainda assim ter baixa Usabilidade. Portanto, o MIS sendo um mtodo que avalia apenas a comunicao, em especial, apenas a emisso da comunicao, considero que o esforo dispendido para aplic-lo no justificado."*

It is noteworthy that some issues mentioned by the participants were not specific to SIM, but rather inherent to any inspection-based method. For instance, issues related to the fact that SIM does not involve users in the evaluation. Therefore, we did not include these issues in the Tables 1 and 2.

Comparative Analysis Although the questionnaire mentioned advantages and disadvantages of SIM specifically, one of the optional open questions asked them how they compared SIM to other methods they had applied. Table 3 summarizes the comparisons made by the participants. The column *"Advantage or Disadvantage?"* shows if the item corresponds to an advantage or disadvantage regarding SIM. We also present in the column "#" the number of participants that mentioned it.

Regarding the advantages and disadvantages presented in Table 3, done in a comparative way, we can see that some of them (**[C2]**, **[C4]** and **[C9]**) are once again general advantages of inspection-based methods when compared to methods involving user observation. All of them compare SIM, which is an inspection-based method, with CEM, which involves observing users in a controlled environment. As mentioned by [24] a comparison of inspection-based methods with user-based methods must be made with care because they have different purposes. Besides, it is expected that user-based methods require more time, effort and resources than inspection-based methods since they need an infrastructure and users to carry out the evaluations.

The comparison made more frequently by participants is that SIM has a higher cost than other methods **[C1]**. Some participants did not mention which methods they were comparing SIM to, but one mentioned that SIM has a higher cost than Heuristic Evaluation and Usability Testing. This particular participant (P2) had some experience with the three methods (i.e., applied Heuristic Evaluation three times and applied SIM and Usability Testing twice). A comparison between SIM and Heuristic Evaluation, as discussed before, should take into account that they evaluate different properties. However, since there is an overlap of problems that can be identified with both properties [2], it could be an interesting direction for future work. We would expect SIM to have a higher learning and cost application, but it would be interesting to investigate the nature of problems identified and how they inform a possible redesign of the system [26]. This could be useful in giving a better understanding of cost and benefit of SIM, since most HCI researchers and practitioners have a good knowledge and understanding of HE.

The biggest advantage of SIM compared to other evaluation methods is related to its results **[C3]**. The participants think that SIM generates better results than CEM and Heuristic Evaluation, for instance:

"I found SIM more complete than the heuristic evaluation because it makes a

deeper analysis of the relationship between system (as deputy) and user."[9](P10)

SIM has also been recognized to generate better results when applied to knowledge generation in the HCI field **[C5]**. Again, there were contradictions among evaluators perception of how SIM compared to other methods. Directly contradicting **[C3]** some participants mentioned as a disadvantage that SIM generates worse results than CEM **[C6]**.

In addition, SIM was considered a difficult method compared to other methods **[C7]**, to require more theory knowledge **[C12]** and not as appropriate for technical evaluations as Heuristic Evaluation **[C8]**. On the other hand, SIM was considered to have a better cost-benefit **[C10]** and to require fewer resources than other evaluation methods **[C11]**. The participants who mentioned advantages/disadvantages **[C10]**, **[C11]** and **[C12]** did not mentioned which methods they were comparing SIM to.

It is important to note that in most of the comparisons made by the participants they did not mention the method they were comparing SIM to. Therefore, it was very hard to generate conclusions about SIM in relation to other evaluation methods. Besides, there is no way of knowing whether participants had any data to support their statements, or whether these statements resulted from feelings they had based on their experiences. Therefore, these issues raised in comparing the methods are not taken as final results about the method, but rather as aspects that could be interesting to further investigate.

We believed that the comparisons would be based on an informal perspective and experience of each evaluator. However, if there had been a consensus it could have been an important indicator. The inconsistencies generated may be related to both the experience of the evaluators and the contexts in which the methods were applied, and indicates that it would be important to make formal assessments to compare the methods according to relevant criteria.

4.2 The Interview

The interview was conducted with four authors of SIM. They will be referenced as I1, I2, I3 and I4. All the authors interviewed have used several HCI methods. They all have experience with Semiotic Engineering methods, as expected, and at least one usability evaluation method. All of them have experience with Heuristic Evaluation, but some have also experience with Usability Testing and Cognitive Walkthrough. Their overall experience in the HCI field ranges from 9 to 20 years.

The interviews were conducted through an instant message system (i.e., Skype, MSN Messenger or GTalk) in December 2011 and lasted, on average, 1 hour and 30 minutes.

[9]Original statement in Portuguese by P10: *"Achei o MIS mais completo do que a avaliao heurstica por fazer uma anlise mais profunda da relao sistema (como preposto) e usurio."*

Table 3. SIM advantages and disadvantages in comparison to other evaluation methods.

	Comparison	Advantage or Disadvantage	Participants	#
C1	SIM has a higher cost than other methods (e.g., Heuristic Evaluation and Usability Test).	Disadvantage	P2; P11; P15; P20; P24; P25	6
C2	SIM has a lower cost compared to CEM.	Advantage	P4; P11; P17	3
C3	SIM generates better results than CEM and Heuristic Evaluation.	Advantage	P4; P10; P25	3
C4	SIM is easier and simpler to apply than CEM.	Advantage	P3; P19	2
C5	SIM generates better results when applied to knowledge generation.	Advantage	P7; P9	2
C6	SIM generates worse results (i.e., less detailed and effective) than CEM.	Disadvantage	P19; P22	2
C7	SIM is more difficult compared to other evaluation methods (e.g., Heuristic Evaluation).	Disadvantage	P3; P9	2
C8	SIM is not a good alternative for technical evaluations compared to Heuristic Evaluation.	Disadvantage	P7	1
C9	SIM is less appropriate than CEM when we want information from user's perspective.	Disadvantage	P10	1
C10	SIM has a better cost-benefit than other methods.	Advantage	P24	1
C11	SIM requires fewer resources than other evaluation methods.	Advantage	P22	1
C12	SIM requires more theory knowledge than other methods.	Disadvantage	P16	1

4.2.1 Results and Analysis The analysis of the interview aimed at investigating the advantages, disadvantages, costs and benefits of SIM from the authors' perspective, which represents the expert evaluators' view. In addition, some issues identified during the survey with novice evaluators were also placed on the agenda during the interview.

SIM's Applicability The applicability of SIM to a wide range of domains was already expected by the authors of the method who raised this hypothesis when they formalized the method [5]. All the authors also mentioned in the interview to have applied SIM to different domains, for instance, collaborative systems, educational systems and text editors. This feature was also reinforced by I2 who said that *"the cool thing about SIM is that it is not dependent on the domain"*[10].

In addition, the authors emphasized that to apply SIM in the domains mentioned there was no need to adapt the method. According to I1, *"adjustments generally fall into a common characteristic of interpretive methods, which is how the evaluator 'fits' the concepts and materials which he/she will work with"*[11]. One may argue that this could be considered an adaptation of the method, however, the same author complements mentioning that the step of framing the concepts and materials is always required when applying SIM: *"... before any application of SIM the evaluator must look at the situation at hand and ask himself/herself how he/she will interpret the static, dynamic and metalinguistic signs"*[12]. It is important to mention that this expectation and experience of the authors related to SIM's applicability is consistent with the previous study described in [6] that showed that SIM can be applied to different domains and technologies without adaptations.

Another point mentioned by the authors of the method is the possibility to investigate, besides communicability, other properties. In other words, SIM allows evaluators to expand their analysis beyond the scope of communicability, allowing the identification of breakdowns related to other properties. According to I1, it is always possible to identify issues related to other qualities of use *"... because communication is the support process of other processes (cognitive, productive, etc.). Thus, when faced with communication problems (or even with certain characteristics of communication, which is not itself a PROBLEM) we always end up anticipating or just taking a glimpse of issues related to usability, productivity, etc."*[13]. Note that capital letters were included by the participant.

[10]Original statement in Portuguese by I2: *"O barato do MIS que ele no dependente de domnio."*
[11]Original statement in Portuguese by I1: *"As adaptaes em geral recaem na parte caracterstica dos mtodos interpretativos, que como o avaliador 'enquadra' os conceitos e materiais com os quais ele vai trabalhar."*
[12]Original statement in Portuguese by I1: *"Antes de qualquer aplicao do MIS o avaliador tem de olhar para o caso em questo e se perguntar como ele vai interpretar o que so signos estticos, dinmicos e metalingusticos."*
[13]Original statement in Portuguese by I1: *"... pois a comunicao o processo de sustentao de outros processos (cognitivos, produtivos, etc.). Ento, ao nos depararmos com problemas de comunicao (ou mesmo com certas caractersticas da comunicao, que no so propriamente um PROBLEMA) sempre acabamos antevendo ou entrevendo problemas de usabilidade, de produtividade, etc."*

While most of the authors agreed with this feature, author I2 believes that this characteristic should not be highlighted. He/she argues that the method should be good at doing what it is proposed to do. For instance, using SIM to find usability problems may end up frustrating the evaluator, because SIM is sure to be weaker to identify usability problems than other methods developed with focus on evaluating usability. Although I2 recommends this caution, I1 makes it clear (in the above excerpt) that the fact that SIM evaluates the communication process makes it possible to identify problems related to other properties because the communication also supports other processes. Therefore, allowing other properties to be investigated is certainly a feature and can be a benefit in certain application contexts. However, the evaluator, when selecting SIM as a method, should take into account that the focus of SIM is evaluating communicability and consider if it would be an appropriate method if other properties are of interest.

Regarding the possibility of identifying domain-specific issues (i.e., if applied to a collaborative system, could it identify specific issues regarding users' interaction through the system, or only general problems, such as static elements that were not consistent with the behavior associated to them?) most of the authors think it is possible to identify domain-specific issues and mentioned to have identified issues specific to the domain in their evaluations with SIM. However, I1 mentioned that this characteristic was never the goal in the application of SIM and he/she prefers to be cautious in confirming this feature of the method. The author believes that a responsible answer would depend on having more knowledge than in fact he/she has. At the same time he/she mentioned to be working on a scientific application of SIM and thinks that he/she is coming close to stating a potentially new concept in the domain in which SIM is being applied. But still he/she thinks that it is too early to say.

Evaluators' Experience With regard to the influence of the evaluators experience the authors agree that the evaluator experience has an influence on the results. Author I4 mentions that it is the evaluator's HCI experience that will allow him/her to make associations between communicability problems and other HCI problems. If evaluators do not have HCI experience they might not be able to analyze what the problems identified in the signs mean regarding interaction.

Author I3 went further mentioning that in addition to HCI experience the evaluator's skills and abilities that come from life experience also influence the results. *"The ability to criticize, analyze and be attuned to the context of what is happening, I think all this ends up influencing the results of SIM application"*[14].

Although the authors agree that the evaluator's experience influence on SIM's application, some of them argue that this does not differ from other existing methods as mentioned

[14]Original statement in Portuguese by I3: *"A capacidade crtica, analtica e estar antenado ao contexto do que est acontecendo, acho que isso tudo acaba influenciando nos resultados da aplicao do MIS."*

by I1: *"... in any evaluation method – and PARTICULARLY in inspection-based methods – the evaluators experience is everything"*[15]. For instance, I4 mentioned that he/she works with Heuristic Evaluation and *"students also find it difficult to assign a heuristic to a problem identified"*[16].

But at the same time author I3 believes that *"SIM, in a way, by being a deeply interpretive method, ends up instigating this more than others"*[17]. According to I2 the Heuristic Evaluation, for instance, gives an illusion that it is easy, fast and simple, but the HCI experience is an implicit prerequisite of the method. SIM is different because while the other methods leave the necessity of expertise in HCI implicit, SIM, because it is based on theory and has chained steps, makes it clear and explicit that the experience is essential. In other words, the fact that SIM is a deeply interpretive method makes it depend more on the evaluators' experience than other methods.

In addition, the importance of the theory knowledge was emphasized by the authors. Author I2 mentioned that the elaboration of a broader and competent opinion about the meta-communication depends on how much the evaluator knows the theory. *"If you do not know [the theory], you will be able to do it, but will be more at the lowest level of abstraction"*[18]. That is, the more one knows about Semiotic Engineering the richer the analysis of the results will be. According to I4, *"students who have a superficial knowledge about Semiotic Engineering (which is the reality of the teaching context) can apply the method, but the results are superficial. They cannot analyze the consequences of problems"*[19].

Costs/Disadvantages and Benefits/Advantages We also asked the authors what were the costs/disadvantages and benefits/advantages of SIM. Regarding costs and disadvantages of SIM the authors I1, I2 and I3 mentioned that SIM requires evaluators to have skills and abilities that are not trivial. According to I1, SIM requires the steps of segmented analysis to be followed strictly as proposed by the method. He/she believes, for instance, that it is very difficult to be strict and not include dynamic signs in the static signs analysis step. Also it requires accuracy in the consolidation of the segmented analysis into a single diagnosis about the communicability of the meta-message. Therefore, to perform an application of

[15]Original statement in Portuguese by I1: *"Em qualquer mtodo de avaliao – e PARTICULARMENTE nos mtodos de inspeo – a experincia do avaliador tudo."*
[16]Original statement in Portuguese by I4: *"os alunos tambm sentem dificuldades para atribuir uma heurstica a um problema que identificam."*
[17]Original statement in Portuguese by I2: *"O MIS, de certa forma, por ser um mtodo profundamente interpretativo, acaba instigando mais isto."*
[18]Original statement in Portuguese by I2: *"Se voc no souber, ser capaz de fazer, mas ficar mais no nvel mais baixo de abstrao."*
[19]Original statement in Portuguese by I2: *"Alunos que tem um conhecimento superficial sobre Engenharia Semitica (que a grande realidade do contexto de ensino) at conseguem aplicar o mtodo, mas os resultados so superficiais. Eles no conseguem analisar as consequncias dos problemas."*

SIM the authors believe it requires the evaluator to have a good abstract and interpretative reasoning. As mentioned by I2 and I3 this is a more specific cost of SIM. According to I3, Heuristic Evaluation, for instance, has a recipe to be followed and students can follow and apply it without problems. Whereas, although SIM has well-defined and chained steps, students often have difficulties in applying it; they cannot make the necessary abstractions, and end up applying it in the wrong way.

A disadvantage mentioned by I2 is that SIM does not involve users, but he/she emphasizes that this is a characteristic and disadvantage of all inspection-based methods (i.e., it is not an exclusive feature of SIM). The same author also mentioned as a disadvantage the fact that SIM requires a functional prototype to be used. In his/her view there should have a formative version of SIM, in other words, that supports the evaluation of interfaces early in the process.

In addition, I3 mentioned as a disadvantage the fact that SIM does not have a support tool. This drawback may be associated with what was mentioned by I4 that applying SIM demands time. The authors believe that this effort to apply it is not necessarily a disadvantage of the method. They argue that the effort associated to its application is related to a cost that should be considered when choosing the method to evaluate a system. This characteristic of SIM (i.e., that it demands time) is justified by the fact that it is a qualitative method, because *"... since it is a qualitative method, in order to have richer results more time is needed."*[20] (I4).

Although the authors agree that SIM demands time, all of them mentioned a great advantaged of SIM that offsets this cost: SIM allows obtaining richer and deeper results. According to I4, SIM allows the evaluator to obtain *"... a thorough understanding of the system to be evaluated."*[21]. Author I1's view is in line with this statement and he/she goes further:

> *"The main advantage (which in my opinion beats all the costs) is to be able to learn a wealth of opportunities, resources and facets of human-computer interaction. Nobody comes out of an application of SIM without learning lots of things about HCI. Not even me. :) That's why I think that SIM is the most valuable tool of Semiotic Engineering today. As Don Norman, himself, has already drawn attention in more than one of his articles, the great advantage of Semiotic Engineering, which exemplary crystallizes in SIM, is the integrated view of how many and so many things that are involved in an HCI project. This contrasts very strongly with other methods that have a more sporadic or fragmented, or*

[20] Original statement in Portuguese by I4: *"Como um mtodo qualitativo, para que o avaliador tenha resultados mais ricos necessrio mais tempo."*
[21] Original statement in Portuguese by I4: *"...um conhecimento aprofundado sobre o sistema a ser avaliado."*

superficial performance."[22]

5 Discussion

We could observe that there was a consensus among novices and authors for most of the SIM characteristics raised. They perceive as the costs of SIM the high time/effort demanded, which can be related to the fact that the method is deeply qualitative and interpretive, and, therefore, demands time and effort to obtain richer results. In addition, the knowledge and skills needed to use it were also mentioned as a cost. First, there is a cost associated to learning the method and its underlying theory. According to I4, *"the cost is high because you need to understand well Semiotic Engineering."*[23]. Semiotic Engineering knowledge is seen as essential for the successful application of the method, and, therefore, it is a cost to be considered when choosing the method. Secondly, there are costs related to skills and abilities that SIM requires of the evaluator that are not easily taught, as for example, *"... how to think, reflect, interpret and abstract"* [24].

The difficulty related to learning the method has also been identified in the work of [29] in which the author notes that the cause of these difficulties could be considered due to the need to change a widely used way of thinking in the computer science field, which is more predictive, accurate and repeatable. She also pointed out that these difficulties are not exclusive of SIM and CEM. Interpreting, abstracting and building a global vision has been identified as the cause for serious difficulties in the teaching and usage of other methods, for example, in programming, interface design and usability engineering. Thus, we conclude that SIM does require an effort to be learned due exclusively to the fact of being a theory-based method and requiring its knowledge to apply it. In addition, the application of SIM requires certain skills that are not easily taught and, probably, learned.

Regarding the effectiveness, SIM was considered by novice evaluators as an effective method since it allows identifying, in most cases, relevant problems and formalizing perceived problems. In the interview the authors complemented mentioning that SIM allows the evaluator to obtain a thorough understanding of the system being evaluated, generating richer and detailed results about the system's communicability. The authors also pointed out that

[22]Original statement in Portuguese by I1: *"A principal vantagem (que a meu ver bate todos os custos) APRENDER a riqueza de oportunidades, de recursos e de facetas da interao humano-computador. Ningum sai de uma aplicao do MIS sem ter aprendido MONTES de coisas sobre IHC. Nem eu. ;) Por isto, acho que o MIS a ferramenta mais valiosa da EngSem atualmente. Como o Don Norman, mesmo, j chamou a ateno em mais de um artigo dele, a grande vantagem da EngSem, que se cristaliza exemplarmente no MIS, a viso integradora de quantas e tantas coisas esto envolvidas no projeto de IHC. Isto contrasta muito fortemente com outros mtodos que tm uma atuao mais pontual ou fragmentria, ou superficial."*

[23]Original statement in Portuguese by I4: *"O custo alto pois preciso compreender bem a Engenharia Semitica."*

[24]Original statement in Portuguese by I3: *"... ensinar a pensar, refletir, interpretar e abstrair."*

this is the main advantage of SIM and that it beats all the costs. According to I2, *"it is costly, but worthwhile"* [25].

Novices and authors agree that SIM allows a thorough analysis of the system, supporting the evaluator in formalizing, arguing and explaining the problems encountered. It also produces rich and deep results about the system's communicability and has a good cost-effective ratio. In addition, the novice evaluators' perception about SIM's applicability is aligned with the authors view. The advantage related to the fact that the method can be applied to different domains without adaptations was mentioned by the novice evaluators and reinforced by the authors. The fact that the method allows identifying the impact of problems related to other system properties (e.g., accessibility, sociability) was also a consensus.

Being a theory-based method was contradictory among novice evaluators (some mentioned as a disadvantage the fact that it is based on a complex theory, whereas others perceived it as a positive aspect). We argue that the disadvantage aspect could probably be related to the cost mentioned previously (i.e., that SIM is a difficult method with a high learning curve). Again this reinforces the cost regarding knowledge and skills necessary to apply the method. On the other hand, one author strengthened the view that being a theory-based method is an advantage of SIM. According to I1, *"methods based on theory are naturally more powerful in the sense that their results talk to concepts that are beyond the specific situation or context of the application"* [26]. Author I2 reinforces this view stating that the theory *"brings results integrated and consolidated into a chained reasoning"*[27]. Although some novice evaluators have mentioned the theory as an advantage, some of them did not include any advantages related to the theory when asked about the advantages of the method. As shown, the authors of SIM have a clearer view of the benefits related to its theoretical basis. This clearer view can be justified by the fact that not only do they have a larger experience with the method and Semiotic Engineering, but also with HCI in general.

Regarding the disadvantages, the lack of tools to support the application of the method was mentioned by both novices and authors. Although mentioned as a disadvantage, we do not characterize as a drawback because this can be easily solved by creating a support tool. Given the results obtained in this study, we believe that the effort is worth it and is already being developed in our research group [30].

Finally, it is interesting to note that both the novice evaluators and the authors of the method have a similar view in relation to the costs and benefits of MIS. It is recognized by the authors and perceived by novices that SIM demands time to apply and learn it, and requires certain knowledge and skills of the evaluators. Despite these costs, SIM is perceived as a

[25] Original statement in Portuguese by I2: *" custoso, mas vantajoso."*
[26] Original statement in Portuguese by I1: *"Mtodos baseado em teoria so naturalmente mais potentes no sentido de que seus resultados conversam com conceitos que esto alm da situao ou contexto especfico da aplicao."*
[27] Original statement in Portuguese by I2: *"Traz resultados integrados e consolidados em um raciocnio encadeado."*

method that allows evaluators to obtain richer and more detailed results about the system's interface. Furthermore, we may argue that the cost of learning the theory and method is fixed, and not associated to every application of the method. Thus, SIM is perceived as a method with good cost-benefit ratio.

As discussed, we have identified some advantages, disadvantages, costs and benefits of the method, which can be useful to assist not only in deeper investigations of SIM, but also to identify ways to improve SIM's use in the HCI field. The reason of having performed this step by obtaining information from the evaluators' perspective is due to the fact that the evaluators are the ones responsible for deciding which method to use for an interactive system evaluation. Therefore, understanding the costs and benefits from the evaluators' perspective may be important in defining (or revising) a strategy to present and/or teach the method, and also in deciding which deeper assessments of the method would be interesting to conduct. In particular, the disadvantage mentioned by the evaluators related to the lack of educational material points to the need of developing support materials so that the method can potentially be better understood and more widely used.

The assessment presented in this paper has some limitations. As already mentioned, the main limitation is the number of respondents: only 25 participants and 4 experts (i.e., the authors) answered, respectively, the questionnaire and interview. This number is not statistically significant considering the international HCI community. However, we considered a good number of responses since SIM is a relatively recent method, and it is not widely used or taught (yet).

In addition, the qualitative analysis of the participants' responses and generation of the advantages and disadvantages lists were done only by the author of this research, which could be a threat to the validity of the results. However, to avoid bias in the analysis the results were discussed with the advisor and a review process was also conducted.

6 Final Remarks

This paper presented an assessment of SIM from evaluators' perspective. We applied a questionnaire to novice evaluators and interviewed the authors of the method. The analysis performed was of a qualitative nature, as opposed to focusing in quantifying the associated cost of applying the method, as has been done in other works (e.g., [19]). Rather the goal was to outline what is perceived as advantages, disadvantages, costs and benefits from the (novices and experts) evaluators' perspective, and also points the differences and similarities between their views.

We raised interesting insights and characteristics of the method, which can be useful to assist not only in deeper investigations of SIM, but also to identify ways to improve SIM's

use in the HCI field. Evaluators are responsible for deciding which method they will use when evaluating an interactive system. Thus gathering information about how SIM is being perceived by the evaluators is important to assess the consolidation of the method. In addition, identifying the advantages and disadvantages from the evaluators' perspective may be important in defining (or revising) a strategy to present and/or teach the method. In particular, the disadvantage mentioned by the evaluators related to the lack of educational material (confirming what was identified by [29]) points to the need of developing support materials so that the method can potentially be better understood and more widely used.

Moreover, understanding the costs and benefits of the method allows a better assessment of the method's potential and definition of new research directions. For instance, the fact that it demands a high cost of learning and application could mean that perhaps we should focus on the use of the method as a tool for scientific evaluations and not for technical evaluations. On the other hand, the fact that novice evaluators find that, despite the cost, it has an associated benefit may be promising for its use as a technical evaluation tool. Thus, this work helps to point future works that can be performed. Some future works include, but are not limited to: (1) analyzing whether the initial cost to learn the method decays over time – this would require to monitor evaluators using the method over a longer period of time; and (2) analyzing whether a system to support SIM evaluation would improve the application cost of the method, and if so, in which aspects.

The results presented in this paper also contribute to the SemEng research, since it generates data regarding how a theory-based method is being used and perceived. This contribution is relevant not only for SemEng research, but also to the HCI field as a whole, since it has already been identified the need to research new methods [3], as well as HCI theories and methods based on them [4, 2].

7 Acknowledgements

Soraia thanks CAPES and Raquel thanks INCT-Web (INWeb) (MCT/CNPq/ grant 57.3871/2008-6) for the support to their research. Authors also thank all participants of the assessment conducted.

References

[1] J. Preece, Y. Rogers, and H. Sharp, *Interaction Design: Beyond Human-Computer Interaction*. New York: Wiley, 2 ed., 2007.

[2] C. S. de Souza, *The Semiotic Engineering of Human-Computer Interaction*. MIT Press, 2005.

[3] S. Greenberg and B. Buxton, "Usability evaluation considered harmful (some of the time)," in *Proceedings of the twenty-sixth annual SIGCHI conference on Human factors in computing systems*, CHI '08, (New York, NY, USA), pp. 111–120, ACM, 2008.

[4] J. M. Carroll, *HCI models, theories, and frameworks: toward a multidisciplinary science*. Morgan Kaufmann, 1 ed., 2003.

[5] C. S. de Souza, C. F. Leito, R. O. Prates, and E. J. da Silva, "The semiotic inspection method," in *Proceedings of VII Brazilian symposium on Human factors in computing systems*, IHC '06, (New York, NY, USA), pp. 148–157, ACM, 2006.

[6] S. d. S. Reis and R. O. Prates, "Applicability of the semiotic inspection method: a systematic literature review," in *Proceedings of the Brazilian Symposium on Human Factors in Computing Systems*, IHC '11, (New York, NY, USA), pp. 177–186, ACM, 2011.

[7] L. F. H. Bento, R. O. Prates, and L. Chaimowicz, "Using semiotic inspection method to evaluate a human-robot interface," in *Proceedings of the 2009 Latin American Web Congress*, LA-WEB '09, (Washington, DC, USA), pp. 77–84, IEEE Computer Society, 2009.

[8] B. A. M. Mattos, R. L. Santos, and R. O. Prates, "Investigating the applicability of the semiotic inspection method to collaborative systems," in *Simpsio Brasileiro de Sistemas Colaborativos (SBSC)*, pp. 53 –60, oct. 2009.

[9] E. R. d. Oliveira, "Investigao sobre a aplicabilidade dos mtodos de avaliao de comunicabilidade ao domnio educacional," Master's thesis, MSc Dissertation in Computer Science - Universidade Federal de Minas Gerais, 2010. 206p.

[10] K. Hornbk, "Dogmas in the assessment of usability evaluation methods," *Behav. Inf. Technol.*, vol. 29, pp. 97–111, January 2010.

[11] C. S. Peirce, *The essential Peirce*. Indiana University Press, 1992.

[12] C. S. de Souza and C. F. Leito, *Semiotic engineering methods for scientific research in HCI*. Morgan & Claypool, 2009.

[13] C. S. de Souza, C. F. Leito, R. O. Prates, S. A. Bim, and E. J. da Silva, "Can inspection methods generate valid new knowledge in hci? the case of semiotic inspection," *International Journal Human-Computer Studies*, vol. 68, pp. 22–40, January 2010.

[14] R. O. Prates, C. S. de Souza, and S. D. J. Barbosa, "Methods and tools: a method for evaluating the communicability of user interfaces," *interactions*, vol. 7, pp. 31–38, January 2000.

[15] R. H. Hartson, T. S. Andre, and R. C. Williges, "Criteria for evaluating usability evaluation methods," *International Journal of Human-Computer Interaction*, vol. 15, no. 1, pp. 145–181, 2003.

[16] C.-M. Karat, R. Campbell, and T. Fiegel, "Comparison of empirical testing and walkthrough methods in user interface evaluation," in *Proceedings of the SIGCHI conference on Human factors in computing systems*, CHI '92, (New York, NY, USA), pp. 397–404, ACM, 1992.

[17] J. Nielsen, "Finding usability problems through heuristic evaluation," in *Proceedings of the SIGCHI conference on Human factors in computing systems*, CHI '92, (New York, NY, USA), pp. 373–380, ACM, 1992.

[18] J. Nielsen and R. Molich, "Heuristic evaluation of user interfaces," in *Proceedings of the SIGCHI conference on Human factors in computing systems: Empowering people*, CHI '90, (New York, NY, USA), pp. 249–256, ACM, 1990.

[19] R. Jeffries, J. R. Miller, C. Wharton, and K. Uyeda, "User interface evaluation in the real world: a comparison of four techniques," in *Proceedings of the SIGCHI conference on Human factors in computing systems: Reaching through technology*, CHI '91, (New York, NY, USA), pp. 119–124, ACM, 1991.

[20] H. Desurvire, J. Kondziela, and M. E. Atwood, "What is gained and lost when using methods other than empirical testing," in *Posters and short talks of the 1992 SIGCHI conference on Human factors in computing systems*, CHI '92, (New York, NY, USA), pp. 125–126, ACM, 1992.

[21] H. Desurvire, D. Lawrence, and M. Atwood, "Empiricism versus judgement: Comparing user interface evaluation methods on a new telephone-based interface," *SIGCHI Bull.*, vol. 23, pp. 58–59, October 1991.

[22] M. P. Steves, E. Morse, C. Gutwin, and S. Greenberg, "A comparison of usage evaluation and inspection methods for assessing groupware usability," in *Proceedings of the 2001 International ACM SIGGROUP Conference on Supporting Group Work*, GROUP '01, (New York, NY, USA), pp. 125–134, ACM, 2001.

[23] P.-Y. Y. Yen and S. Bakken, "A comparison of usability evaluation methods: heuristic evaluation versus end-user think-aloud protocol - an example from a web-based communication tool for nurse scheduling.," in *Proceedings of the Symposium AMIA*, pp. 714–718, 2009.

[24] W. D. Gray and M. C. Salzman, "Damaged merchandise? a review of experiments that compare usability evaluation methods," *Human-Computer Interaction*, vol. 13, pp. 203–261, Sept. 1998.

[25] G. Lindgaard, "Notions of thoroughness, efficiency, and validity: Are they valid in hci practice?," *International Journal of Industrial Ergonomics*, vol. 36, no. 12, pp. 1069–1074, 2006.

[26] L. C. d. C. Salgado, S. A. Bim, and C. S. de Souza, "Comparao entre os mtodos de avaliao de base cognitiva e semitica," in *Proceedings of VII Brazilian symposium on Human factors in computing systems*, IHC '06, (New York, NY, USA), pp. 158–167, ACM, 2006.

[27] P. Coopamootoo and D. Ashenden, "A systematic evaluation of the communicability of online privacy mechanisms with respect to communication privacy management," in *Design, User Experience, and Usability. Theory, Methods, Tools and Practice* (A. Marcus, ed.), vol. 6770 of *Lecture Notes in Computer Science*, pp. 384–393, Springer Berlin / Heidelberg, 2011.

[28] A. Strauss and J. Corbin, *Basics of qualitative research: Techniques and procedures for developing grounded theory*. SAGE, 1998.

[29] S. A. Bim, *Obstculos ao ensino dos mtodos de avaliao da Engenharia Semitica*. PhD thesis, Ph.D Thesis on Informatics - Pontifcia Universidade Catlica do Rio de Janeiro, 2009. 181p.

[30] J. S. Dias and R. O. Prates, "Sistema de apoio aplicao do mis," in *Companion Proceedings of the 11th Brazilian Symposium on Human Factors in Computing Systems*, IHC '12, (Porto Alegre, Brazil, Brazil), pp. 35–36, Brazilian Computer Society, 2012.