Muhammad Nazrul Islam

How Does User Interpretation of Interface Signs Affect Web Usability

An Empirical Case Study of a Web Application

Turku Centre for Computer Science

TUCS Technical Report
No 1000, February 2011

How Does User Interpretation of Web Interface Signs Affect Web Usability: An Empirical Case Study of a Web application

Muhammad Nazrul Islam

Åbo Akademi University, Department of Information Technologies,
Joukahainengatan 3-5 A, FI- 20520 Turku, Finland

TUCS Technical Report
No 1000, February 2011

Abstract

In the present day User Interfaces (UIs) are complicated software components, which play a crucial role in the usability of web applications. An explosion on interface design for HCI has been commenced over the last decade. But very little attention has been paid to semiotics theories for web interface design, though designing the web sign has a widely acceptable crucial effect on enhancing users understanding and satisfaction. For these, the objective of this paper is to reflect user experiences in interface signs interpretation and how these could affect the usability of web applications. To accomplish this objective, a systematic empirical case study was conducted on a web application. This study was replicated with seven participants from five different educational institutions in Finland and followed a strict case study methodology to ensure the validity and reliability of our research outcomes. This paper presented the case study design and discussed the results achieved. Finally, this paper also talks about future trends of semiotics theory as applied to interface design and evaluation.

Keywords: think aloud usability testing, semiotics, interface signs, web application, user experiences

1. Introduction

Over the past years, with the advent of globalization and the rise of information technology, it has become obvious that one of the most important qualities of web application would be the ease by which the end user can learn and interact with these applications. Thus, the activity of assessing the quality degree of the applications is becoming an arduous task. Users' degree of satisfaction in using as well as interacting with a web application established the quality of this application [1]. The most significant measurement unit of satisfaction is usability, as it is “the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use” [20]. Thus, usability of a system contributes to user happiness, satisfaction, as well as pleasure; and conversely a lack of this contributes to user dissatisfaction and frustration, and thus eventually will result in the total abandonment of the system. Therefore, usability is considered a key quality for a web application.

The web interface plays the main role for the interaction between human and computer in web applications. The growing demand of the present Internet world leads us to focus on designing these web interfaces, user perception on web interface signs, web usability as well as the crucial roles of designing web interface signs to HCI. In fact, these design principles are semiotics by nature and semiotics is the science of signs [2], that is, a theory about sense production and interpretation. For this, by exploring semiotic perceptions to interface signs, new and important perspectives for interface design will be discovered.

Therefore, the purpose of this research was to discuss semiotics theories and to show the significance of semiotics to design and also to evaluate the web interface signs to boost web usability. Indeed, this research shows that semiotics is one of the important fundamental design dimensions that affect the usability of a web application. In this research a systematic empirical case study on a web application was conducted and revealed how user's interpretation to interpret interface signs could affect the overall usability concerns.

This paper is structured as follows. In section 2, the previous research related to our work is described. Semiotics theories and these relations to interface sign interpretation are discussed in section 3. These semiotics theories were also the sources of motivation for us to design and articulate this paradigm. This section also highlights the spectrum of issues that were encounter in user intuitive test. The steps of experimental method for empirical study are discussed in section 4. An empirical case study on web application is discussed in section 5. In section 6, our research findings are presented. Finally, the conclusion concerning semiotics perspective to interface sign design as well as ideas of future research is provided in section 7.

2. State of the Art and Related Works

Over the last few decades, usability evaluation method has been considered as an important quality assessment technique in website evaluations [3]. Thus, different Usability Evaluation Methods (UEM) have emerged and been developed in research and practice in the field of usability. UEM can be divided into four classes: analytic methods, specialist reports (usability inspection), observational methods (usability testing) and user reports (survey) [4]. Analytic methods are mainly driven by analysis [5] of tasks that need to be done by the end users. Usability inspections focus on feedback from experts in HCI or web application design. Heuristic evaluation [6] [7], cognitive walkthroughs [8], feature inspections [7] are more common evaluation methods to this group. Checklists, usability principals or rules are used as guidelines to direct this kind of evaluations. Co-discovery [9], think-aloud [8] [10] are effective example methods of usability testing group. In co-discovery, two or more users work together in the evaluation. For the think-aloud, a small number of users are involved individually, users verbalize while using the system to complete the given task to express his/her thoughts, feelings, and opinions. The final group, user report involves the use of questionnaires [10] [11] and interviews [10] [11] for data collection.

All the evaluation methodologies presented above are lacking the evaluation of semiotic issues of web applications. These methods do not analyze the intrinsic values of user interface, specially the interface signs of user interface. To allow the analysis of intrinsic values of interface signs during usability evaluation, a semiotic engineering approach has been evolved [12]. However, current well-structured web usability evaluation methods and techniques consider semiotic aspects as generic criteria for evaluating the user satisfaction, often confusing and blending them with other usability problems (i.e. problems related to content or to layout design) [1]. Moreover, very few methods give the right importance of semiotic design and evaluation to optimize the web usability.

The main reasons to skip semiotics issues in the currently available UEM of web applications as well as for designing the interface signs are: (i) lack of knowledge on semiotics and its theories in general, (ii) lack of theoretical background on semiotics theories to web interface sign design and its evaluation, (iii) lack of understanding the necessities of semiotics to interface design and evaluation, (iv) lack of awareness on how semiotically designed interface signs affect the web usability, etc. This research has mainly focused on these issues and shows how users’ understanding of interface signs affects web usability, and thus eventually presents the significance of semiotics theories to design and evaluate the interface signs through a systematic empirical case study on a web application.

3. Semiotics and Sign Interpretation

Among the many different semiotics models two models are presented here which were more relevant to this research work: (i) Peirce's semiotics model [2] consists of a triadic relationship containing: the *representamen (representation or sign)* - this stands to somebody for something in some respect or capacity. It addresses somebody and creates in the mind of that person an equivalent, or perhaps more developed sign; the *object (referent)* - is the actual thing the sign stands for and the *interpretant (meaning)* - is therefore the sign created in the mind of the perceiver or the reaction caused by the object in the perceiver [13]. For these, a sign requires the concurrent presence of these three constituents. Let us give an example: consider a panel at the entrance of a company with “Reception office” written on it. The textual shape of the sign (the text string “Reception office”, the font used, its color, its background, its size, etc.) is the sign. The concept that the sign evokes in the mind of the reader, that is, the idea of a *reception’s office* and what it means is the interpretant. The actual object in the real world, that is, the reception’s office as physical object is the referent.

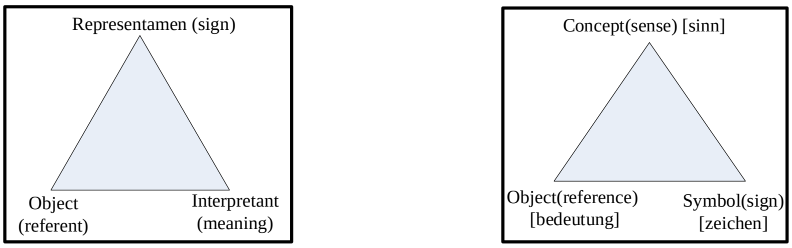


Figure 1. Peirce’s semiotic triangle (left side) and Frege’s semiotic triangle (right side)

(ii) Semiotics theory by Gottlob Frege's terms for the three vertices of the semiotic triangle were *Zeichen* (sign) for the symbol, *Sinn* (sense) for the concept, and *Bedeutung* (reference) for the object [14]. As an example of the semiotics triangle, Frege cited the terms ‘*morning star’* and ‘*evening star’* and both terms refer to the planet Venus as their meaning, but their senses are very unlike the way in which the planet is presented (one term refers to a star seen in the morning, and other one refers to a star seen in the evening). Therefore, there is no one-to-one link between the object and the sign; various signs may have a single meaning in spite of several meanings. Different signs vehicles can refer to the same object since each sign vehicle has its own flavour or sense that leads it to the same object.

From the above discussion on semiotic theory as well as semiotic model these were found that generally, users guess the sign meanings through the *creation and interpretation* of 'signs'. Signs take the form of words, images, sounds, odours, flavours, acts or objects, but these things have no intrinsic as well as intended meaning and these things become signs only when designers provide these with meaning (or, sense) [15]. An example of Frege’s semiotics theory depicted in figure 2, observes that different signs may lead to the same object by different interpretants. Pierce’s model depicted in figure 3, observes that the same sign may refer to different objects by different interpretants.

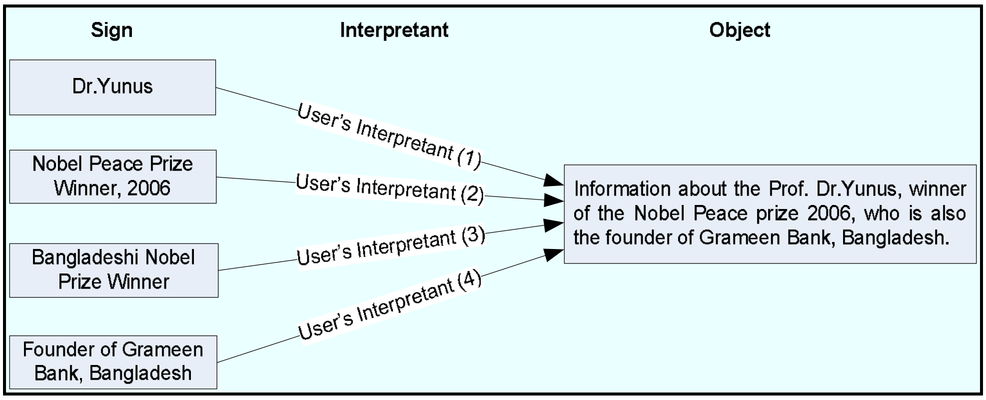


Figure 2. Different signs referring to the same object

For these, the user interpretation (an example is depicted in figure 2) of interface signs were classified into the following categories based on the accuracy level of user interpretation with respect to the designer's interpretation for an interface sign: *a) accurate*- user's interpretation completely matches the designer's interpretation and this category reflects the semiotics theory, *(b) moderate*- user's felt more than one distinct object, one of which was the right one about the interface signs and probability to obtain the right object at the first attempt may be less than the accurate interpretation (for example, if a user proceeds with a sign to obtain a particular object but the sign does not really stands for that), *(c) conflicting*- user's felt more than one distinct object in his/her mind about the interface signs and user felt confused about choosing the right object that will match to the designers intention, *(d) erroneous*- user's interpretation referred to a completely different object other than the designer's interpretation, and *(e) incapable*- user could not able to interpret the interface sign at all. These categorizations were also used in empirical studies in section 5.

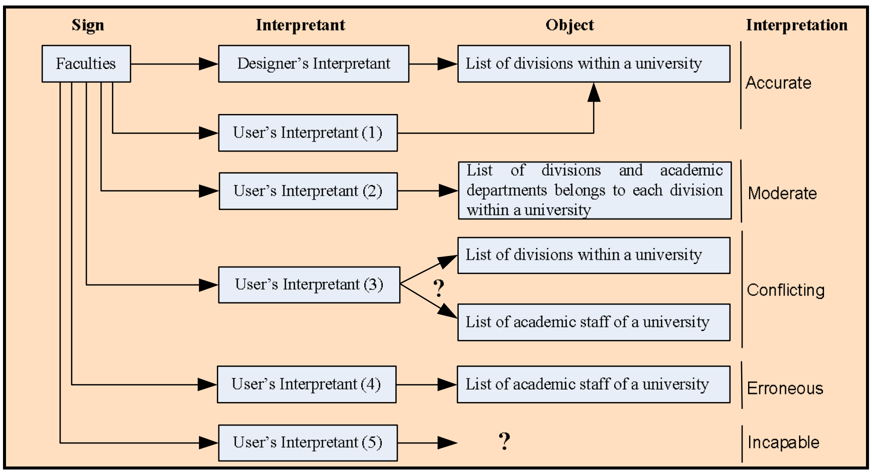


Figure 3. Possible Interpretation of a web interface sign

4. Experimental Method

In this research, two user tests were conducted (i) User intuitive test and (ii) Conventional think-aloud usability test to reach at research goal. To obtain the research outcome through these tests the seven sequential steps were followed (see figure 4). In the following sections, these steps are discussed in more details with an example case study. Briefly these are the seven steps.

Step 1: The problem statement and test objectives were clearly defined to reflect the purposes of conducting the tests and appropriately derive the remaining steps.

Step 2: Tasks list were prepared and then all the interface signs were listed along with listing (separately) the entire related interface signs (heuristically) to these tasks for the web application being tested.

Step 3 & 4: Participants who might be the user of this studied application were recruited and scheduled. The laboratory was set up to conduct these tests properly.

Step 5: A user intuitive test were conducted to understand the user interpretations of these listed interface signs and collected the data in a systematic way.

Step 6: User testing to do the given tasks was conducted following the conventional laboratory based think-aloud method and collected the data in a systematic way.

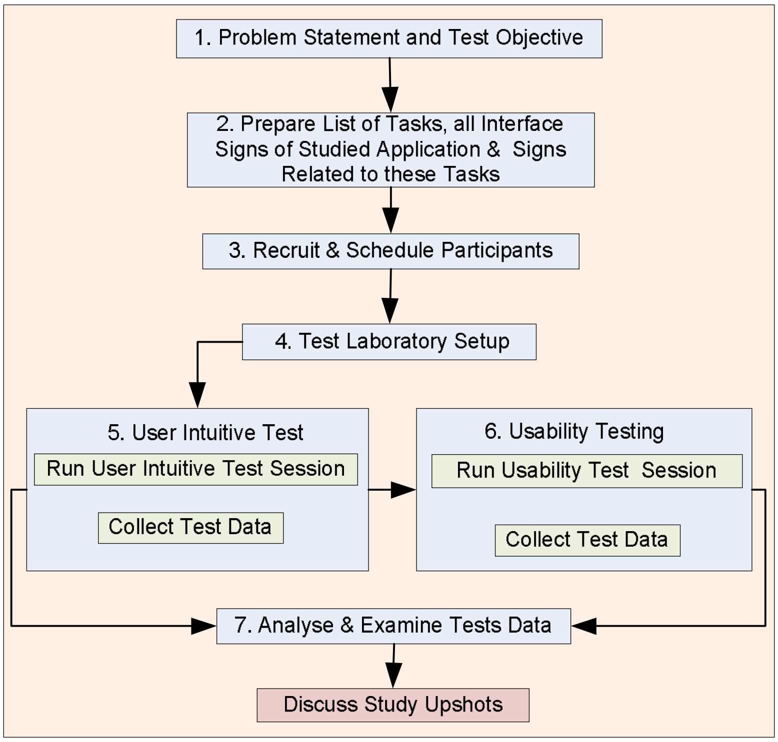


Figure 4. Structure of our experimental methodology

Step 7: Finally, these tests data were analyzed and examined to observe the user behaviour focusing the users' understanding of interface signs and how these understandings affect users' performance. For example, an interface sign S is related to a task T. From the user intuitive test, if it was happened that a user U does not understand properly the intended meaning of S. Then, it was observed from the usability testing data (video clips, facilitator's notes, think aloud verbalization etc.) - (a) how the user U behaves to do the task T while the sign is S related to this task T was not properly understandable to U. (b) how these behaviours influence the web usability.

The studies ended by discussing the important observations that emerged from the analysis phase, and also presented the future trends of semiotics theory as applied to interface design and evaluation.

5. An Empirical Study on Web Application

The purpose of our empirical study following the above systematic procedure was to observe the user experiences to deal with interface signs and how these could affect the web usability. This study has been conducted on an example web application, the Ovi calendar of Ovi by Nokia (http://calendar.ovi.com) during the period of week 47 to 49, 2010 at usability testing laboratory of Åbo Akademi University, Finland. This section briefly discusses how the empirical study was conducted to reach at our research goal.

5.1. Problem Statement and Test Objective

This study mainly focuses on user understanding of interface signs of the web application from a semiotics perspective. One basic research question is addressed:

How do interface signs (semiotics) affect web usability?

Our test objective was to obtain the answer of this research question. In particular, this research wanted to observe:

- user understanding (accuracy level) to interpret the intended meaning of interface signs.
- user behaviour to do a specific task with respect to his/her understanding (accuracy level) to this task-related interface signs.
- And how their behaviour influenced web usability.

5.2. Tasks and Interface Signs

A set of scenarios were created where each scenario contained multiple tasks. The scenarios were written in the language of user's tasks. The scenarios and its related tasks are briefly presented in table 1.

After finalizing the tasks list, an inspection (heuristically) was carried out very meticulously to find (i) all interface signs, and (ii) possible related interface signs to each task. Then two lists of interface signs were prepared: one having all interface signs of the Ovi calendar and another having related interface signs to each of these tasks.

Table 1. List of scenarios and related tasks

Scenarios	Task no.	Tasks
Log-in & event entry	1	log-in to Ovi calendar (data was provided)
	2	create an event
	3	create an event with advanced options
Search & edit event	4	search for an event (event entered previously)
	5	edit an event
check, delete and log-out	6	check weekly event list
	7	delete an event
	8	log-out to leave Ovi calendar

5.3. Participants

Anyone who wants a personal, free calendar service that can be accessed from any location from any web browser might be the users of this product. This would mean that the people, who have web access from home, work, or even from other remote locations would be the users. Therefore, students were chosen as our test users. Due to limitations of time and money, this study did not cover other types of users. A series of questionnaires were designed to qualify the potential users. The overall study involved seven male participants aged 21 - 30, selected from five different universities (Åbo Akademi University, University of Turku, University of Tampere, Turku University of Applied Sciences, Novia University of Applied Sciences) in Finland. All participants had good experience in using the personal computer, the internet, the real world calendar and three users had prior experience in using a web calendar, but no participant had prior experience in using the chosen application (Ovi calendar).

Table 2. Test participants profile in brief (H: High, M: Medium, L: Low, N: None)

Features	P1	P2	P3	P4	P5	P6	P7
Familiarity with personal computer	H	H	H	H	H	H	H
Internet familiarity	H	H	H	H	H	H	H
Age	21	22	22	26	25	29	24
Education	M	M	M	H	M	H	H
Familiarity with real world calendar	H	H	H	H	H	H	H
Familiarity with online Ovi calendar	N	N	N	N	N	N	N
Familiarity with other online calendar	L	N	H	N	N	H	N

5.4. Test Laboratory setup

The usability laboratory had two rooms, a test evaluation room and an observation room. The usability evaluation room was furnished with a video camera to record the videos of the users' activities, a computer on the desk running on Windows Operating

System. For web browsing Internet Explorer, Mozilla, and Google Chrome were available. The computer was connected to the internet through an Ethernet connection. The observation room was furnished with audio/video mixer, recorder, microphone, speaker, computer running Windows Operating System, and Observer 5.0 software. Both tests (user intuitive test and think-aloud usability test) were conducted in usability laboratory. Both test sessions were made recorded on video.

5.5. User Intuitive Test

The user intuitive test was conducted through user interviewing mainly. The main reasons for choosing interviewing were [16]: interviews need very few facilities, easy to organize, enjoyable, as well as a good way to find in-depth information about users. The user interview was conducted one by one following the thinking aloud method [17]. At the beginning, the interviewer gave a very short lecture to the participant regarding the purpose of the interview and web application being tested. The questionnaires used to conduct this test session were: *What do you think about the intended meaning of this sign? / What could be the purpose of using this sign? / What is your guess about the referential content for this sign? / Why does this sign stand for?* The author as interviewer and a participant, as interviewee were seated together in front of a computer, showed the list of all interface signs of web application being tested. Selected participants were asked to formally interpret these interface signs (base on the questionnaires raised for this test session), talking aloud and described their understanding of each interface signs. Page snapshots of studied application from where signs were listed were also showed and they were asked to “re-comment” on any signs if they thought their past comments were not appropriate to any particular sign. The fundamental purpose of asking these questions was to obtain an indication of their understanding and classify their interpretation into: accurate, moderate, conflicting, erroneous and incapable (see section 3).

During the test session, the interviewer was careful of the two main things: to prompt the user to keep up the flow of comments and provide help when necessary. But the interviewer was always alert while prompting and helping them during interview to avoid distorting the results. For example, if someone missed any important interface element, then a word from interviewer helps them to focus their attention right on it. Since research shows that people will make up an answer to any question when asked, whether or not they have any basis for the answer. Moreover, users were requested to avoid the duplicate signs for the successive pages, if they already commented on those before. The interviewer noted these data during test sessions and these entries were checked again with the video record of the test sessions.

5.6. Think-Aloud Usability Test

The discount usability testing approach was followed to test the Ovi calendar. The approach involves testing only a small number of participants to yield reliable results, thus limit cost and time. Research shows that carefully designed usability testing with 5 participants might reveal 80% of the usability problems [18]. Even though discount usability evaluation is not guaranteed to find a large portion of the usability problems [19], in this research, where observing user behaviour in interacting with interface signs was the scope of evaluation , it was sufficient to conduct with 7 participants.

To perform a usability test with each participant, the following activities were followed. A short lecture was given about the system and usability testing in general. Activities during test sessions consisted of observing users performing their tasks in a usability test laboratory. Their activities were recorded in videos and they were observed through a one way mirror. Post-task questionnaires were used to obtain immediate feedback of the users after completing each scenario. The users were asked about the ease and difficulty of tasks and provided options to write comments on different issues they felt during the completion of tasks. It helped to obtain feedback when users' memory was fresh. At the end, when users finished last scenario, post-test questionnaires were delivered. Later the video record of the test sessions were examined and coded using data-logging software (Observer 5.0) to obtain test data. Apart from this, more data were collected from different kinds of questionnaires, for example pre-test, post-task, post-test.

5.7. Analysis and Examine the Tests Data

Data from both tests were collected to analysis into two steps: (i) general analysis, and (ii) critical analysis. Microsoft Excel 2007, Spotfire Decision Site 7.3, and Observer 5.0 software were used to analysis and examine these data.

5.7.1. Data Collection

From user intuitive test, data of - (i) users' interpretations of all 104 interfaces signs (see table 3) as well as (ii) users' interpretations of each task-related signs (see table 4) were collected.

Table 3. Categorizations of participants' interpretations to all interface signs

Participants	Accurate	Inaccurate				Accuracy (%)	Inaccuracy (%)
		Moderate	Conflict	Erroneous	Incapable		
P1	67	18	3	12	4	64,42	35,58
P2	65	13	8	16	2	62,50	37,50
P3	79	16	0	8	1	75,96	24,04
P4	73	14	5	4	8	70,19	29,81
P5	71	14	3	11	5	68,27	31,73
P6	77	11	2	9	5	74,04	25,96
P7	76	9	4	13	2	73,08	26,92

Table 4. Users' interpretations of task-related interface signs

Participants	Task	Related signs	Accurate	Inaccurate				Accuracy (%)	Inaccuracy (%)
				Moderate	Conflict	Erroneous	Incapable		
P1	1	4	4	0	0	0	0	100	0
	2	48	26	13	1	7	1	54,17	45,83
	3	48	26	13	1	7	1	54,17	45,83
	4	17	7	5	0	5	0	41,18	58,82
	5	52	30	13	1	7	1	57,69	42,31
	6	12	6	1	0	5	0	50	50
	7	21	10	5	0	5	1	47,62	52,38
	8	1	1	0	0	0	0	100	0
P2	1	4	4	0	0	0	0	100	0
	2	48	29	11	2	6	0	60,42	39,58
	3	48	29	11	2	6	0	60,42	39,58
	4	17	8	3	0	6	0	47,06	52,94
	5	52	32	11	2	7	0	61,54	38,46
	6	12	5	1	0	6	0	41,67	58,33
	7	21	11	3	1	6	0	52,38	47,62
	8	1	1	0	0	0	0	100	0
P3	1	4	4	0	0	0	0	100	0
	2	48	33	8	0	6	1	68,75	31,25
	3	48	33	8	0	6	1	68,75	31,25
	4	17	9	4	0	3	1	52,94	47,06
	5	52	36	12	2	7	0	69,23	30,77
	6	12	5	2	0	4	1	41,67	58,33
	7	21	12	4	0	4	1	57,14	42,86
	8	1	1	0	0	0	0	100	0
P4	1	4	4	0	0	0	0	100	0
	2	48	33	7	1	1	6	68,75	31,25
	3	48	33	7	1	1	6	68,75	31,25
	4	17	9	5	0	3	0	52,94	47,06
	5	52	36	7	2	1	6	69,23	30,77
	6	12	5	5	0	0	2	41,67	58,33
	7	21	13	5	0	3	0	61,90	38,10
	8	1	1	0	0	0	0	100	0
P5	1	4	4	0	0	0	0	100	0
	2	48	33	9	0	3	3	68,75	31,25
	3	48	33	9	0	3	3	68,75	31,25
	4	17	8	4	1	2	2	47,06	52,94
	5	52	36	9	1	3	3	69,23	30,77
	6	12	4	4	1	1	2	33,33	66,67

	7	21	11	4	1	3	2	52,38	47,62
	8	1	1	0	0	0	0	100	0
P6	1	4	4	0	0	0	0	100	0
	2	48	34	5	0	7	2	70,83	29,17
	3	48	34	5	0	7	2	70,83	29,17
	4	17	11	2	0	4	0	64,71	35,29
	5	52	37	5	0	7	3	71,15	28,85
	6	12	7	2	0	3	0	58,33	41,67
	7	21	15	2	0	4	0	71,43	28,57
	8	1	1	0	0	0	0	100	0
P7	1	4	4	0	0	0	0	100	0
	2	48	35	6	1	5	1	72,92	27,08
	3	48	35	6	1	5	1	72,92	27,08
	4	17	9	3	1	4	0	52,94	47,06
	5	52	38	6	2	5	1	73,08	26,92
	6	12	5	3	1	3	0	41,67	58,33
	7	21	10	4	1	6	0	47,62	52,38
	8	1	1	0	0	0	0	100	0

From think-aloud usability test, data of – (i) task completion time (TCT), min and max time for each as well as all tasks (see table 5); (ii) number of times tried / failed to complete each task (see table 6); (iii) number of input error, system error as well as number of times despaired, smile, angry, asking help for each task (see table 7); (iv) number of interaction and interact variation for each task (see table 8); (v) time to stay at despaired, smiley, or angry state (see table 9); (vi) time to stay at confused & wrong navigation (C&WN) state (see table 10); (vii) subjective rate in the scale of 1-5 based on how easy or difficultly felt to do each task as well as overall reaction to the studied application (see table 11); and (viii) examples of verbal comments related to interface sign interpretation when they think out loud (see table 12) were collected for all participants.

Table 5. Task completion time (mm:ss); min and max time cell are coloured as light turquoise and rose respectively

Participants	T1	T2	T3	T4	T5	T6	T7	T8	Total
P1	01:28	05:44	08:26	00:44	05:40	02:19	01:10	00:07	25:38
P2	01:00	04:02	18:53	00:55	04:55	02:02	02:59	00:03	34:49
P3	00:38	02:57	06:47	01:44	15:18	00:29	00:36	00:03	28:32
P4	01:15	03:08	02:54	00:59	03:16	05:35	00:44	00:05	17:56
P5	01:24	06:23	07:43	00:54	01:19	06:25	01:01	00:03	25:12
P6	01:03	07:00	02:21	01:35	11:40	05:08	03:22	00:03	32:12
P7	01:12	03:11	14:07	03:04	21:27	04:39	00:14	00:03	47:57

Table 6. Number of times tried (and failed i.e., *no. of times tried -1*) to complete the tasks

Participants	T1	T2	T3	T4	T5	T6	T7	T8
P1	1	1	2	1	1	1	1	1
P2	1	1	2	1	1	1	2	1
P3	1	1	2	1	3	1	1	1
P4	1	1	1	1	1	2	1	1
P5	1	1	1	1	1	2	1	1
P6	1	1	1	1	2	1	1	1
P7	1	1	1	1	3	1	1	1

Table 7: Number of asking helps (A), input error (I), system error (S), despaired (D), smile (Sm), angry (Ag)

Participants	T1	T2	T3	T4	T5	T6	T7	T8
P1	-	Sm	A,A,I,S	-	A	A	A	-
P2	-	-	A,A,A,A,D,D,D	-	A	D	A	-
P3	-	-	S,A,Sm	-	A,A,D,D,D,D,Ag	-	-	-
P4	-	-	S	-	A	D,D	-	-
P5	-	A	D,A,D	-	-	D,A	-	-
P6	-	A	-	-	A,D,Ag	D,A,Ag	-	-
P7	-	-	D	-	D,D,D,D,A,A,Ag	A,D	-	-

Table 8. Interact variation (*i.e., user interact - required interaction*) to interface signs.

Task	Number of Interaction to Interface Signs		
	Required	User Interact	Variation
T1	3	3	0
T2	12	29	17
T3	27	41	14
T4	3	3	0
T5	15	40	25
T6	2	25	23
T7	2	5	3
T8	1	1	0
Total	65	147	82

P1

Task	Number of Interaction to Interface Signs		
	Required	User Interact	Variation
T1	3	3	0
T2	12	15	3
T3	27	119	92
T4	3	10	7
T5	15	35	20
T6	2	19	17
T7	2	32	30
T8	1	1	0
Total	65	234	169

P2

Task	Number of Interaction to Interface Signs		
	Required	User Interact	Variation
T1	3	3	0
T2	12	22	10
T3	27	48	21
T4	3	35	32
T5	15	183	168
T6	2	6	4
T7	2	8	6
T8	1	1	0
Total	65	306	241

P3

Task	Number of Interaction to Interface Signs		
	Required	User Interact	Variation
T1	3	4	1
T2	12	17	5
T3	27	25	-2
T4	3	17	14
T5	15	55	40
T6	2	55	53
T7	2	7	5
T8	1	1	0
Total	65	181	116

P4

Task	Number of Interaction to Interface Signs		
	Required	User Interact	Variation
T1	3	4	1
T2	12	35	23
T3	27	64	37
T4	3	6	3
T5	15	14	-1
T6	2	68	66
T7	2	28	26
T8	1	1	0
Total	65	220	155

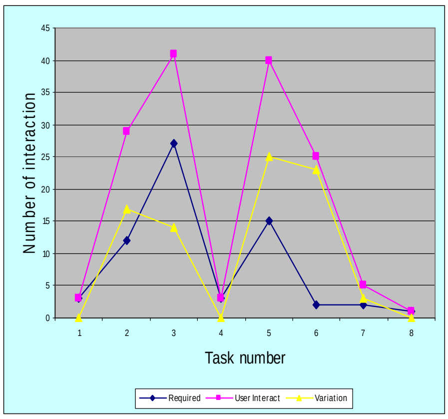
P5

Task	Number of Interaction to Interface Signs		
	Required	User Interact	Variation
T1	3	4	1
T2	12	45	33
T3	27	22	-5
T4	3	17	14
T5	15	122	107
T6	2	37	35
T7	2	40	38
T8	1	1	0
Total	65	288	223

P6

Task	Number of Interaction to Interface Signs		
	Required	User Interact	Variation
T1	3	4	1
T2	12	8	-4
T3	27	126	99
T4	3	35	32
T5	15	243	228
T6	2	57	55
T7	2	3	1
T8	1	1	0
Total	65	477	412

P7



Example Graph of interact variation for P1

Table 9: Facial expression (mm:ss)

Participants	Despaired	Smile	Angry
P1	-	0:16 (T2-0:16)	-
P2	4:42(T3-4:30, T6-0:12)	-	-
P3	4:52 (T5-4:52)	0:04(T3-0:04)	0:22(T5-0:22)
P4	1:58 (T6-1:58)	-	-
P5	3:47 (T3-0:47, T6-3:00)	-	-
P6	5:34 (T5-4:09, T6-1:26)	-	0:34(T5-0:07, T6-0:27)
P7	9:30 (T3-3:11, T5-4:50, T6-1:29)	-	0:12 (T5-0:12)

Table 10: Confused and wrong navigation state (mm:ss)

Participants	Confused and wrong navigation (C&WN) state
P1	4:14 (T2-1:34, T3-0:32, T5-1:03, T6-1:05)
P2	6:08 (T3-4:01, T5-0:29, T6-0:58, T7-0:40)
P3	6:01 (T5-6:01)
P4	4:07 (T3-0:03, T5-0:30, T6-3:34)
P5	7:51 (T2-4:51, T3-0:47, T6-2:13)
P6	10:02 (T2-1:13, T4-0:24, T5-5:14, T6-2:54, T7-0:17)
P7	16:44 (T3-7:39, T4-0:52, T5-6:38,T6-1:35)

Table 11. Subjective rating (1: very difficult/ frustrating , 5: very easy/satisfying)

Participants	T1	T2	T3	T4	T5	T6	T7	T8	Overall satisfaction
P1	5	4	3	4	4	5	5	5	4
P2	5	3	3	5	3	5	5	5	3
P3	5	4	4	3	3	5	5	5	4
P4	5	5	5	5	5	2	5	5	4
P5	5	4	4	5	5	1	2	5	4
P6	5	4	4	3	2	3	4	5	1
P7	5	4	1	1	1	1	1	5	1

Table 12. Examples of verbal comments related to interface sign interpretation

Comments	Task - Participant	Reasons
"Oops!!!"	T3-P3	Wrong interaction with a sign while he knew that the right one is the above one.
"Oh! Everything is lost! Where are my items that I entered? Seems very bad!!"	T4-P3	When he did not find his entered events in the calendar body using day, month, and week view.
"How come I can't repeat? Why repeat option is off?"	T5-P3	When he missed to edit from first date of a continuous event (event with repetition).
"Why I am not finding this(entered event)"	T6-P4	While participants trying to find entered events on the calendar body but he did not obtain this.
"I am not finding 'Delete'. How come I am not getting the 'Delete' option?"	T5-P6	Participants trying to delete already entered event but did not find the 'Delete' option.
"Am I deleted this"	T3-P7	did not understand show delete and retrieving sings. So, from deleted item list he clicked on a deleted event and retrieved this item. So, this deleted item was removed from the delete item list but participant thought that he deleted the actual entered event.
"I have got this! But I didn't want to get this in this way"	T4-P7	Fail to find an event using search option and then while he was trying arbitrarily suddenly this event appear on the calendar body.
"So these are deleted items! I am browsing deleted items!!!"	T5-P7	When participant understood that he was treated the deleted items as the actual entered item list.
"Option is obviously available here but why I am not getting this. Where is the 'Delete' option"	T5-P7	When he was trying to delete an item but after trying comparatively long time he did not obtain this.
"It should come (appear in the calendar body). Trying day, week…"	T5-P7	After editing an event he was trying to see this but he d did not obtain this at calendar body
"All 3 times add this but why I am not getting it"	T5-P7	After editing properly he clicked on save sign but did not obtain the way to check whether is this properly saved or not
"Where should I change options…"	T6-P7	He was sure that there is an entered event in current week but he was trying for comparatively long time but unable to see this on calendar body.

5.7.2. Analysis

5.7.2.1. General Analysis

Taking into account the percentage of users' interpretations accuracy and inaccuracy of all interface sings as well as each task-related signs, general analysis was carried out.

Task completion time: Considering user interpretation of all interface signs this study showed that $P1, P2, P4, P5$ having interpretation accuracy relation $P2 < P1 < P5 < P4$ completed the tasks with task completion time relation $P4 < P5 < P1 < P2$. That is, 57.14% of total participants completed tasks in comparatively shorter time than participants who had comparatively low accuracy of interpreting interface signs. Again, considering task-related interface sign understandability, it was observed that with comparatively high inaccuracy to task-related sings leaded them to do the tasks with comparatively more time {e.g., T2 (sign interpretation inaccuracy 45.83%, task completion time 5:44) by P1, T3 (39.58%, 18:53) by P2, T4 (47.06%, 3:04) by P7, T6 (66.67%, 6:25) by P5, etc.} and vice versa {e.g., T2 (31.25%, 2:57) by P1, T3 (27.08%, 2:21) by P6, T5 (30.77%, 1:19) by P5, etc.}. [Table 3, 4 & 5]

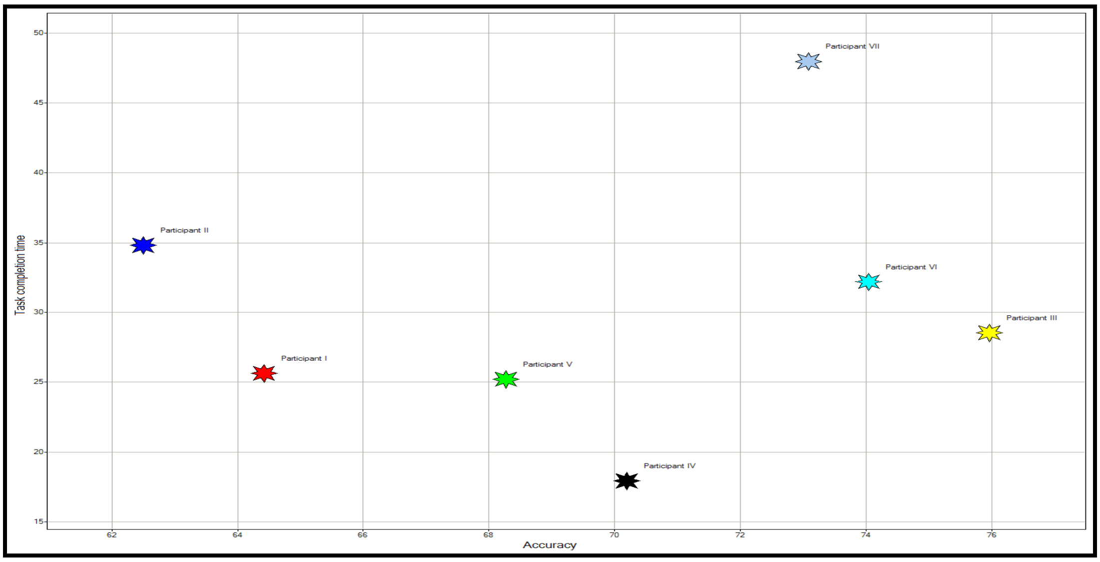


Figure 5. Task completion time versus interpretation accuracy by all participants

Confused & wrong navigation state: In this study, 42.86% of total participants $(P3, P6, P7)$ showed that interpreting intended meaning of interface signs properly spend a comparatively little duration at confusing and wrong navigation than participants who had comparatively low accuracy of sign interpretation thus made comparatively lesser amount of navigation error. Again, considering task-related interface sign understandability, it was observed that with comparative high inaccuracy to task-related sings leaded them to stay at confused & wrong navigation state for relatively more time {e.g., T2 (sign interpretation inaccuracy 39.58%, task completion time 4:01) by P2, T6 (58.33%, 3:34) by P4, etc.} and vice versa {e.g., T3 (27.08%, 0:00) by P6, T3 (31.25%, 0:03) by P4, T1 & T8 (00%, 0:00) by all participants, etc.}. [Table 3, 4 & 10]

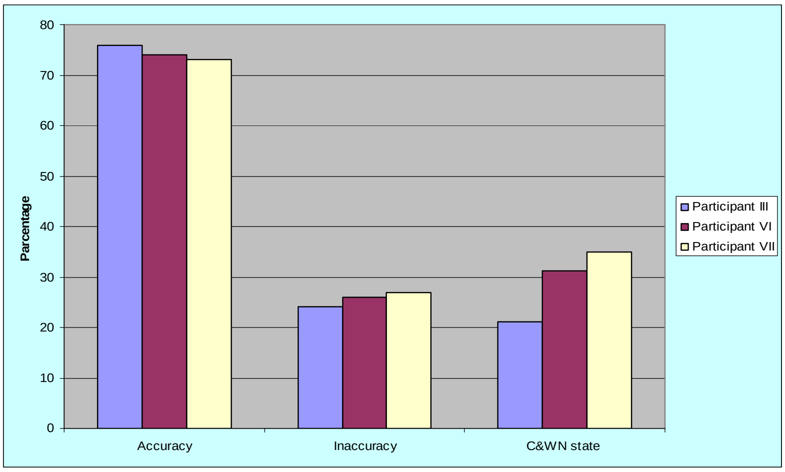


Figure 6. C&WN state with respect to sing accuracy by P3, P6 and P7

Task failure: Total 5 tasks out of 9 (55.56%) were failed by all participants at the first attempt where participants' interpretation of these tasks-related signs showed comparatively higher to inaccuracy. Tests data showed that to do the tasks $T3, T3, T7, T6, T6$ where inaccurate interpretation of interface signs (45.83%, 39.58%, 47.62%, 58.33%, 66.67%) were comparatively higher by participants $P1, P2, P2, P4, P5$ respectively were failed to complete these tasks at the first strive. [Table 4 & 6]

Interact Variation: Most of the cases it was observed that interact variation decreases for a task where participant accurate interpretation was comparatively high for this task-related signs {e.g., T2 (sign interpretation accuracy 68.75%, interact variation 5) by P4, T2 (72.92%, -4) by P7, T8 (100%, 0) by all participants, etc.} and vice versa { e.g., T3 (sign interpretation accuracy 60.42%, interact variation 92) by P2, T6 (33.33%, 66) by P5, etc.}. [Table 4 & 8]

Other error: It was observed that to do task T3 having 45.83% interpretation inaccuracy of this task-related signs by P1, an incorrect interpretation of a sign of date input leaded him to entry an incorrect input (input date in a wrong format since interface signs missing the proper indication of acceptable date format) and this input error generated a system error. Thus, user failed to complete this task at the first attempt and completed this in second attempt. User asked help mostly for tasks, where users' interpretations of interpreting signs were comparatively higher to inaccurate interpretation. For example, to do tasks T3 and T5 where inaccurate sign interpretations were {45.83%, 42.31%} and {39.58%, 38.46%} by PI and P2 respectively were asked help for maximum time. [Table 4 & 7]

Facial expression: In this study, 42.86% of total participants $(P2, P4, P5)$ showed that interpreting intended meaning of interface signs inaccurately spend comparatively more duration at despaired than participants who had comparatively high accuracy of sign interpretation. Considering task-related interface sign understandability, it was observed that with comparative high inaccuracy of task-related sings leaded them to despair as well as angry for relatively more time {e.g., T3 (sign interpretation inaccuracy 39.58%, despaired time 4:01) by P2, T6 (66.67%, 3:00) by P5, T6 (66.67%, 1:26, angry time 0:27) by P6, etc.} and vice versa {e.g., T3 (27.08%, 0:00) by P6, T3 (31.25%, 0:03) by P4, T1 & T8 (00%, 0:00) by all participants, etc.}. [Table 3, 4 & 9]

Subjective ratings: Data from post test questionnaires it was observed that, 57.14% of total participants having comparatively high interpretation accuracy gave subjective ratings of overall satisfaction to comparatively higher than those who had comparatively low interpretation accuracy. For example, P1 had high score of sign interpretation accuracy than P2. Thus P1 gave subjective rate to 4 whereas P2 to 3. [Table 3 & 11]

5.7.2.2. Critical Analysis

Alternate cases to all categories of general analysis were also observed and these cases are named as critical cases. Few examples of critical cases are presents here in below:

- Interpretation accuracy of all interface signs by P4 and P7 were 70.19%, 73.08% respectively but why task completion time, interact variation, confused & wrong navigation, despaired, angry, task failure, asking help, and subjective ratings were comparatively most awful of P7 than P4.

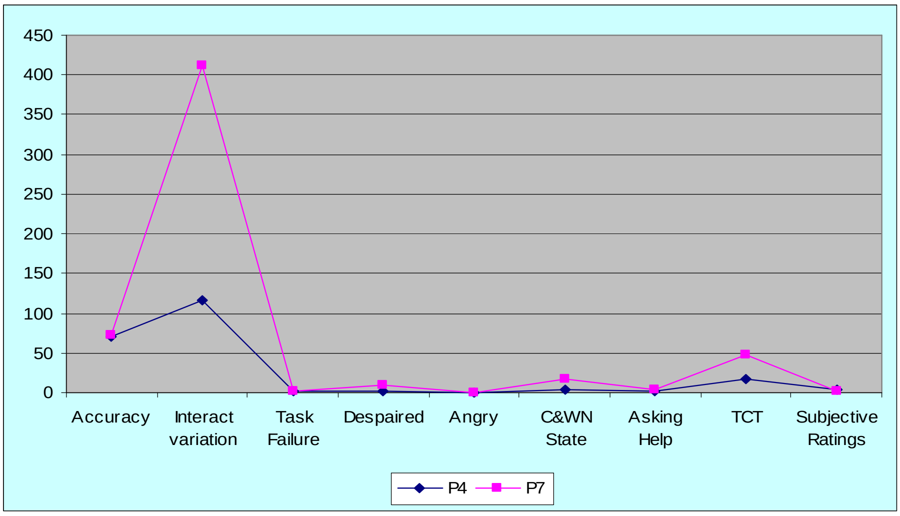


Figure 7. Comparative view between P4 and P7

- To perform the task T5 by P7 having higher accuracy to this task-related interface signs than P6 but why P7 showed comparatively worst performance than P6 on the way to task completion time, interact variation, confused & wrong navigation, despaired , asking help as well as subjective ratings.

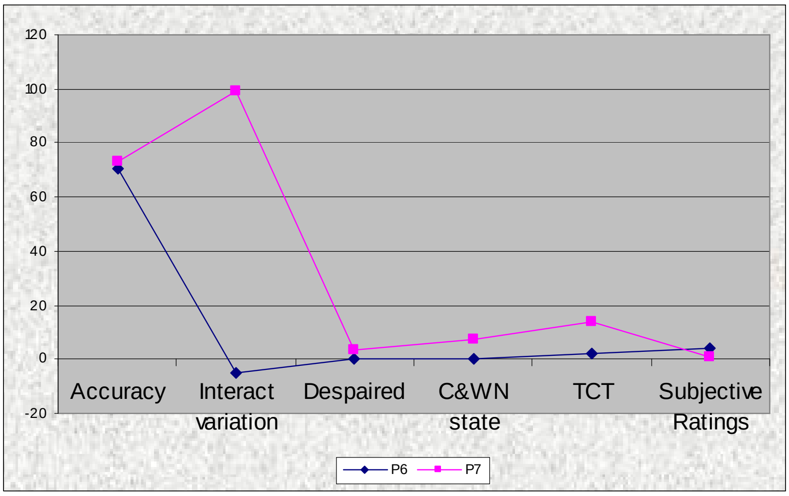


Figure 8. Comparative view between P6 and P7 to perform the task T5

- Inaccurate understandability of task T4 related interface signs by PI and P6 were 58.82% , 35.29% respectively but why interact variation, TCT , C&WN , subjective rating were 0%, 0:44, 0, 4 respectively by P1 as well as 466.67%, 1:35, 0:24, 2 respectively by P6.
- Interface signs for tasks T2 and T3 were showed the same understandability for each participant (e.g., P1's accuracy of interpreting task-related signs for tasks T2 and T3 were 54.17%) but why task completion time, interaction variation, task failure, facial expression, confused and wrong navigates were differed between these tasks by each participant.
- P2 having inaccurate sign interpretation 41.67% for T3 and failed to task completion at the first attempt but with comparatively higher inaccurate sign interpretation (e.g., 46.15%, 44.44% for the task T4 and T6 respectively) didn't failed to complete this (e.g., T4, T6, etc.) task at the first attempt.

Further examined of the tests video leaded to analysis and discuss these critical issues by set operations [21], and dependency graph [22]. Discussions of critical cases are presented here through three observations:

Observation I: It was observed that total number of interface signs of studied application did not exceed the number of any particular task-related signs. Again, except T1 and T8 all other tasks had more than one distinct ways to complete properly and for this, number of required signs to each way of task completion and number of task-related signs was not equal. Therefore, these were happened that participants having comparatively low interpretation accuracy (of all interface signs as well as task-related signs) but understood all required signs for a specific way to do a specific task and proceeded with those signs facilitated him to do that task properly (for example, low task completion time, success at complete task at first attempt, low interact variation etc.). Again, participants having comparatively high interpretation accuracy (of all interface signs as well as task-related signs) but failed to understand all required signs for any specific way (among different distinct ways) to do the specific task showed worst performances for completing this task. This observation is analyzed and discussed here with an example (by set operations) in a more structured way to represent the more clear idea-

Assumptions,

Interface signs for whole system: $S = \{S1, S2, S3, S4, S5, S6, S7, S8, S9, S10, S11, S12, S13, S14\}$

User interprets interface signs accurately for whole system:

$$A = \{S1, S3, S5, S7, S9, S10, S11, S12, S13, S14\}$$

Task: T

No. of distinct ways to perform task T: 4

Required signs for each of these ways:

$$W1:\{S1, S2, S4\}; W2:\{S3, S7, S8\};\ W3:\{S1, S4, S5\}; W4:\{S1, S7\}$$

No. of required signs for each of these ways: $|W1| = 3; |W2| = 3; |W3| = 3; |W4| = 2$

Thus, total *no. of signs for whole system:* $N = |S| = 14$

Interpretation accuracy for whole system: $\{(|A|/|S|) \times 100\} = 71.43\%$

User interprets inaccurately for whole system: $IA = (S - A) = \{S2, S4, S6, S8\} = 28.57\%$

Related signs for task T:

$$RS : (W1 \cup W2 \cup W3 \cup W4) = \{S1, S2, S3, S4, S5, S7, S8\}$$

No. of related interface signs for task T: $|RS| = 7$

For these, $|RS| \le N \quad and \quad |W1| \le |RS| \wedge |W2| \le |RS| \wedge |W3| \le |RS| \wedge |W4| \le |RS|$

That is, no. of task-related signs will be always less than or equal to total number of signs and no. of required signs to each specific way will be always lest than or equal to no. of task-related signs.

User understands accurately for task T: $B = RS \cap A = \{S1, S3, S5, S7\} = 57.14\%$

User understands inaccurately for task T: $C = RS - A = \{S2, S4, S8\} = 42.86\%$

Finally, these assumptions illustrate:

$$RS \subseteq S \wedge (A \cup IA) = S \quad and \quad B \subseteq A \wedge C \not\subset A \wedge B \not\subset IA \wedge C \subseteq IA$$

$$B \subseteq RS \wedge C \subseteq RS \wedge (B \cup C) = RS \quad and \quad W1 \not\subset B \wedge W2 \not\subset B \wedge W3 \not\subset B \wedge W4 \subseteq B$$

According to these assumptions as well as analysis through set operations, though the accurate and inaccurate understanding for this task-related signs were 57.14% and *42.86%* respectively and for all interface signs were 71.43% and 28.57% respectively but if user proceed with *W4* way of task completion then he/she succeed at compete the task T properly and if he/she chose other ways (*W1, W2, W3* or any other arbitrary ways) then task completion results might vary (e.g., increase TCT, fail to complete at the first attempt, interact variation will increase etc.).

Observation II: Further observation as well as examined the tests video showed that to do a specific task in a way (either it has multiple or single way to do this task) required signs were maintained a sequential as well as dependable relations within the set of these required signs. That's why, these were happened that participants having comparatively high accuracy (of all interface signs as well as task-related signs) but incapable or erroneous interpretation of any one (or more) sign(s) among required signs for a specific way to do a specific task showed comparatively most awful performance to task completion time, interact variation, confused & wrong navigation, despaired, angry, task failure, asking help, and subjective ratings than who had comparatively low inaccuracy (of all interface signs as well as task-related signs). For example, to do T3 by P2, T5 by P7 participants did not able to interpret few signs properly within the set of required signs to the specific way chose by them; whereas these few signs were strongly related (sequentially and dependability) with other signs of that set and thus showed worst performance for the specific tasks. This observation is analyzed and discussed here with an example case of doing task T3 by P2 (by dependency graph) in a more structured way to depict the more clear idea. In dependency graph (see figure 9), all circles represent related signs for task T3. Colour represents its understandability to different categories (e.g., red colour represents erroneous interpretation). A duplicate use of sign label S12 refers that same sign were presented twice in studied application. The arrow sign from S2 to S1 means S2 is dependable on sign S1. Among different ways to do this task, circle with labelling represents required signs for a specific way to do a task, in which way P2 was tried to do this task.

Set of required signs for a specific way chose by P2 for T3:

$$\{S1, S2, S3, S4, S5, S6, S7, S8, S9, S10, S11, S12\}$$

Dependency among these signs:

$$\{S1 \leftarrow S2 \leftarrow S3 \leftarrow S4 \leftarrow \{\{S5, S6, S7 \leftarrow S8, S9, S10, S11\} \leftarrow S12\}\}$$

Sequentiality among these signs:

$$\{S1 \Rightarrow S2 \Rightarrow S3 \Rightarrow S4 \Rightarrow \{S5, S6, S7 \Rightarrow S8, S9, S10, S11\} \Rightarrow S12\}$$

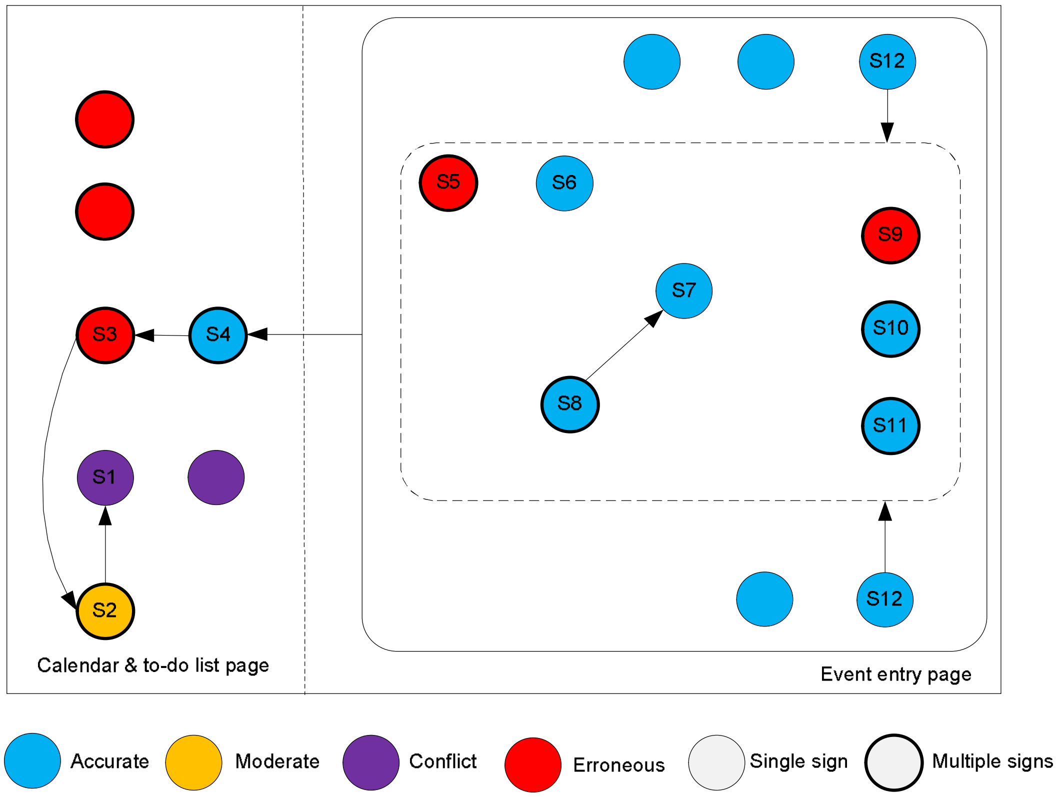


Figure 9. Sign interaction for T3 by P2

That is, S2 depended on S1 (or, to obtain S2 user needed to interact with S1 first), S3 depended on S2 (or, to obtain S3 user needed to interact firstly with S1 then S2), and so on. Dependency as well as sequential relation within a set of required signs for a specific task completion way chose by P2 to do T3 and his interpretation of these signs showed that to reach at event entry page dependable signs (S1, S2, S3) were not properly understandable to him thus he has tried a lot to reach at event entry page. Since, P2 needed to interact with S1 (conflicting) to reach at S2, then needed to interact with S2 (moderated) to reach at S3 and after that needed to interact with S3 (erroneous) to reach at S4 and interaction with this S4 (accurate) leaded him to reach at event entry page to interact with $\{S5, S6, S7, S8, S9, S10, S11, S12\}$. Again, in event entry page erroneous interpretation to S5, and S9 affected to ending sign S12 (sign for saving data of event entry) because S12 was dependable on these signs and sequentially related too and therefore, these leaded him to complete T3 with worst performance (e.g., high TCT, C&WN, task failure, high interact variation etc.) and gave subjective rate for this task to comparatively low score. To do task T5 by P7 same reasons (i.e., incapable to properly interpret few signs) leaded him to show worst performance for task completion and subjective ratings for this task T5.

Observation III: From this study, it was also observed that sign interpretation directly as well as indirectly affected usability matrices of effectiveness (e.g., % of goal achieved), efficiency (e.g., time to complete a task, error rate, amount of effort) and satisfaction (e.g., subjective rating scale) thus eventually affected web usability. For the lack of space only two cases as examples are discussed here, i.e., (i) P1 was unable to properly interpret a sign of input date (in figure 10, sign interpreted by P1) to do task T3 and this inaccurate interpretation made an input error and this input error generated a system error and failed to do this task at the first attempt. Then, these failure and errors showed the way of asking help twice, spend confused & wrong navigation state for 32 second thus make navigation errors and these eventually directed him to increase interact variation (51.85%). Then, these all affected to increase the task completion time (8:26, whereas min time was 2:57) comparatively and finally these affected his subjective rating (rate to 3) too.

(ii) To do task T5 by P7 erroneous and incapable interpretation to few signs (in figure 10, two signs interpreted by P7) leaded him to task failure for first two attempt as well as spend confused and wrong navigation state for 6 min 38 sec (30.91% of task completion time). Then these ultimately directed him to obtain high interact variation (1520%). After that these all affected to increase the task completion time (21:27, whereas min time was 1:19) comparatively and also to his facial expressions (despaired and angry for the time of 4:09 and 0:07 respectively). Finally these affected his subjective rating (rate to 1) too. That is, an erroneous interpretation of interface signs affected directly and indirectly usability matrices (e.g., task failure affects to effectiveness; input error, TCT, interact variation affects to efficiency; asking help, C&WN, subjective ratings to satisfaction) thus eventually affected to overall web usability. Mostly happened possible cases observed from this study are depicted in figure 11. Here, two nodes linked with one side arrow means arrow sided node affected by other side node.

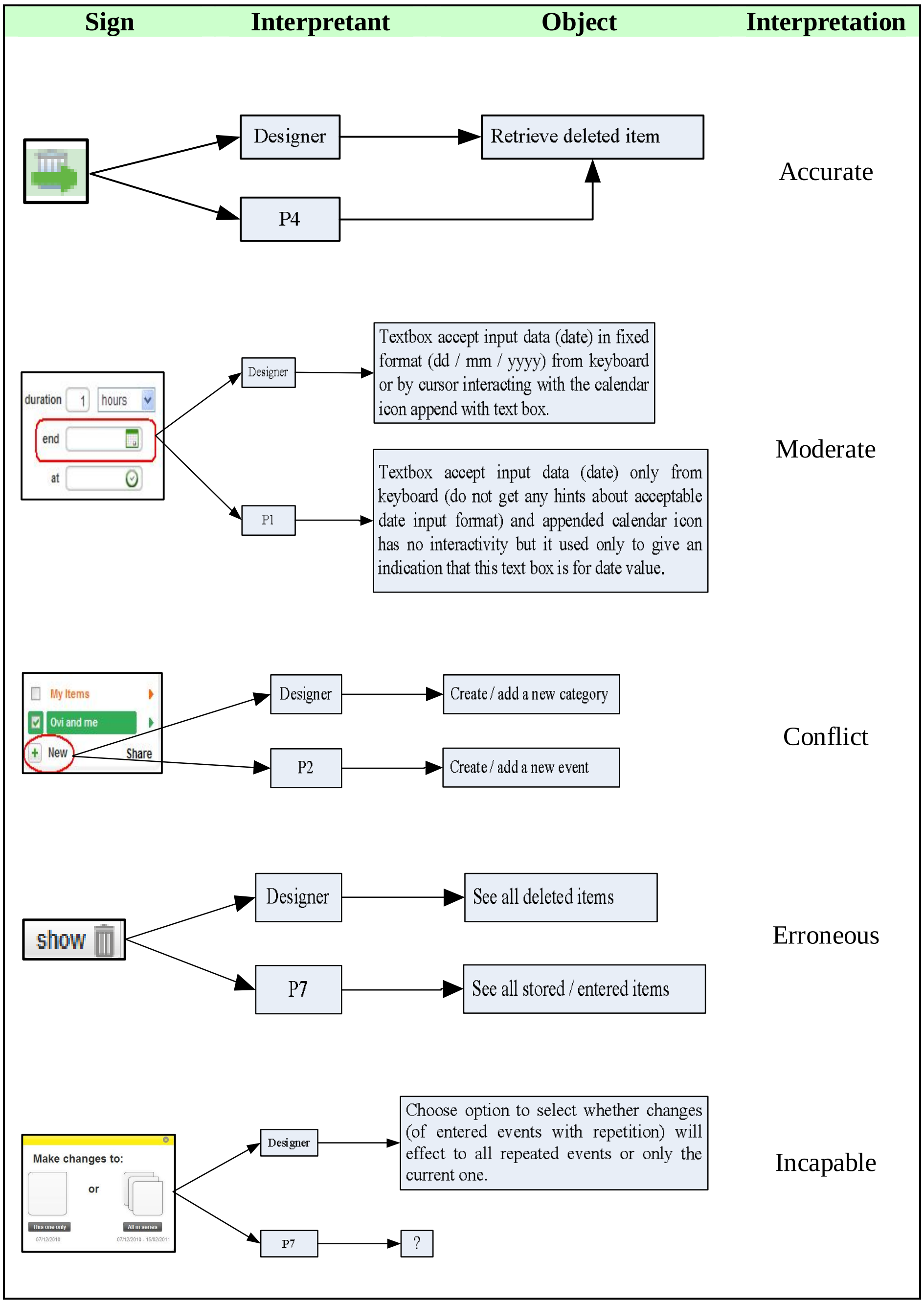


Figure 10. Examples of participants' interpretation of interface signs and its categorization

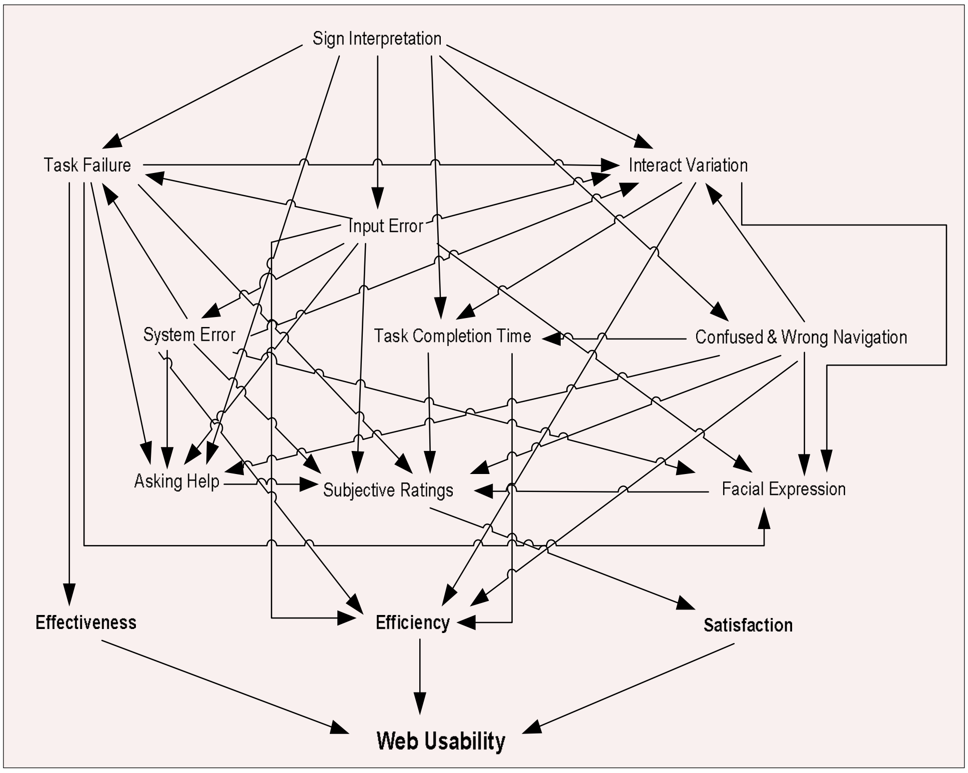


Figure 11. User's interpretation to interface signs affects the web usability

6. Research Findings

Main outcomes of this study are presented here briefly. Users able to interpret interface signs properly complete a task in comparatively shorter time, with lowest interact variation, and spend a shorter time on confusing and wrong navigation than users who do not understand interface signs properly. The possibility to task completion failure rate decreases with proper understanding of interface signs. Erroneous or incapable interpretation of interface signs may lead to task failure. Again, lack of proper interpretation of interface signs related to value input may lead users to make input errors. Interface sign interpretation does not affect the system error directly, but indirectly. It was also observed that ease and ability to interpret the interface signs affects users' facial expressions. Moreover, this study also showed that subjective ratings to overall satisfaction could be comparatively higher to the users who were comparatively more able to interpret interface signs properly.

The number of required signs for a specific way to do a specific task and number of signs related to this task could not be equal. That is, there might be different distinct ways to do a specific task where the number of required signs could also vary but do not exceed the total number of task-related signs. Again, to do a specific task in a specific way (among different distinct ways) the set of required signs could be sequential as well as dependable on other signs within that set of required signs. No matter what is the percentage of interpretation accuracy of all including task-related interface signs, user interpretation might affect task completion performance if - (i) a (or few) sign(s) within the set of required signs for a specific way (chosen by a user) of task completion, and (ii) sequential and dependable relation are also available within these signs. Then critical cases may occur. For example: the user understood all the required signs for a specific way of task completion and proceeded with those signs to do the task properly (e.g., low TCT, low interact variation, success at completing the task at the first attempt etc.), the number of task-related signs and its understandability could be the same for the two distinct tasks for a particular user but their task completion performance might vary in doing these tasks, etc. (i.e., other critical cases).

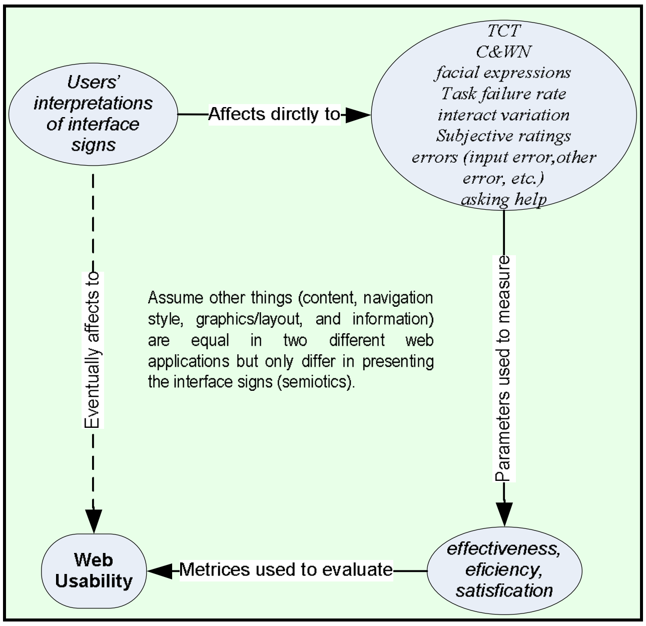


Figure 12. Users' interpretations of interface signs affecting web usability

A web page generally includes: content, navigation, graphics / layout, information, and interface signs. This is why, this study's objective was to depict how interface signs could affect overall usability while considering others were being correct from usability perspective (see figure 11 and 12). This study showed that web interface sign presentation (design) and its interpretations affects most of the problems found through usability test. That is, sign interpretation directly as well as indirectly affects usability matrices (effectiveness, efficiency and satisfaction) thus eventually affecting web usability.

This research did not claim that other things (content, navigation, graphics/layout, Information) are correctly organized in the Ovi calendar interface because this study did not focus on these. The aim of the study was to observe the usability problems in general and these problems were examined with respect to semiotics perspective (users' interpretation of interface sign) to show the importance of considering semiotics acuity to design and evaluate web interfaces to boost web usability.

7. Conclusion and Future Work

This paper shows the importance of semiotics perspective to web usability through an empirical case study on a web application. Firstly, overviews of usability, UEMs, as well as semiotics theories are presented. Then, users' interpretations of interface signs are classified based on the semiotics theories. After these, an example case study is discussed systematically. Finally, the analysis and research findings are presented. This study showed that semiotics consideration to interface design and evaluation were mostly important since interpretation accuracy of interface signs affect usability matrices i.e., effectiveness, efficiency, and satisfaction thus eventually facilitating optimization of the web usability.

There were a few main limitations to this study. Firstly, the case study was conducted only on a web application; secondly, the number of participants were also rather small; thirdly, studied application was tested only on a desktop computer not in a mobile platform; and fourthly, this study did not focus on other contents (e.g., navigation, graphics, etc.) of the web interface therefore there is no evidence of their perfectibility from the usability perspective. The author hopes to consider these issues in future tests.

Again, though many researches have been conducted on web interface, especially on its content, navigational style, graphics/layout, and information, surprisingly web signs were always neglected. Moreover, the answer of a basic research question (*How do interface sign affect web usability?*) accomplish from this research raises another important fundamental question *"What does the designer need to be aware of when re/design meaningful, understandable web interface signs?"* In future work, the author will seek to provide answers for this question. Therefore, this research also acted as an initial step to start my journey to work on a concrete project *"Semiotic Perception in Information Intensive Web Interface: Evaluation and Optimization of Usability and End User Experience"*. Future research on this project will continue by focusing on interface signs, web usability, HCI, UX, and semiotics theories. Considering semiotics perspectives as well as other focused areas, an in-depth empirical study through heuristic evaluation and user testing will be conducted to provide a conceptual interface design and evaluation framework towards optimizing the web usability.

8. Acknowledgements

The author would like to thank all the participants of the study. Åbo Akademi Foundation (ÅAF), Finland provides the grant that has made this research possible. For this, ÅAF is gratefully acknowledged. To Franck Tétard, thank you for your feedback and suggestions for strengthening this paper. The author thanks also Colette Gattoni for her language guidance.

References

[1] Triacca, L. :Wish you were usable! How to improve the quality of a Museum Web site. In: International Conference on Electronic Information, the Visual Arts and Beyond , Florence (2003).

[2] Peirce, C. S.: Collected Writings (8 Vols.).In: Ed. Charles Hartshorne, Paul Weiss & Arthur Burks, Harvard University Press, (1931-58).

[3] Al-Wabil, A., Al-Khalifa, H.: A framework for integrating usability evaluations methods: The Mawhiba web portal case study. In: IEEE International Conference on the Current Trends in Information Technology (CTIT), pp: 1-6, Dubai, UAE (2009).

[4] A. Whitefield, F. Wilson, and J. Dowell,: A framework for human factors evaluation. In: Behavior and Information Technology, vol. 10, pp. 65-79, 1991.

[5] Diaper, D. and Stanton, N.: The handbook of task analysis for human-computer interaction. In: Mahwah, NJ: Lawrence Erlbaum (2003).

[6] Nielsen, J.: Designing Web Usability: The Practice of Simplicity, In: New Riders Press (1999)

[7] Nielsen J. and Mack, R.L.: *Usability inspection methods*. In: Wiley, New York (1994).

[8] Hertzum, M., and Jacobsen, N. E.: The Evaluator Effect: A Chilling Fact about Usability Evaluation Methods. In : International Journal of Human-Computer Interaction (15:1), pp 183-204, (2003).

[9] Barnum, C. M., and Dragga, S. : Usability Testing and Research. In: Longman Publishers, New York (2001).

[10] Nielsen, J.: Usability Engineering. In: Academic Press, Inc., London, UK (1993).

[11] Usability Net: http://www.usabilitynet.org/home.htm. (accessed on Nov 8, 2010).

[12] Souze, D.: Clarisse Sieckenius de Souza: The Semiotic Engineering of Human-Computer Interaction. In: The MIT Press, ISBN: 0262042207, (2005).

[13] Andersen, P.: Computer Semiotics. In: Scandinavian Journal of Information systems, Vol. - 4, pp. 3 -30, (1992).

[14] Frege, G.: Begriffsschrift, English translation. In: J. van Heijenoort, ed., pp. 1-82, Harvard University Press, Cambridge, MA, (1879). .

[15] Morris, C.: Foundations of the Theory of Signs. In: International Encyclopedia of Unified Science, vol. 1, no. 2, University of Chicago Press, Chicago, (1938).

[16] Online resource: http://www.webcredible.co.uk/user-friendly-resources/web-usability/user-interviews.shtml (accessed on Nov 3, 2010).

[17] Lewis, C.: Using the thinking-aloud method in cognitive interface design. In: IBM Research Report RC 9265, Yorktown Heights, NY, (1982).

[18] Nielsen, J.: Guerrilla HCI: Using Discount Usability Engineering to Penetrate the Intimidation Barrier. In: Bias, R. G. and Mayhew, D. S. (eds.), Cost-Justifying Usability, Academic Press, Inc., pp 245-272 , Boston, MA (1994).

[19] Faulkner, L.: Beyond the Five-User Assumption: Benefits of Increased Sample Sizes in Usability Testing. In: *Behaviour Research Methods, Instruments, & Computers* (35:3), pp 379-383, (2003).

[20] ISO 9241 -11:1998, Guidance on usability standards. Available from http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=16883 (accessed on November, 2010).

[21] Set (computer science). Available in: http://en.wikipedia.org/wiki/Set_(computer_science) (accessed on December, 2010).

[22] Dependency graph. Available in: http://en.wikipedia.org/wiki/Dependency_graph (accessed on December, 2010).

Turku Centre *for* Computer Science

Joukahaisenkatu 3-5 B, 20520 Turku, Finland | www.tucs.fi

University of Turku

- Department of Information Technology
- Department of Mathematics

Åbo Akademi University

- Department of Information Technologies

Turku School of Economics

- Institute of Information Systems Sciences

ISBN 978-952-12-2555-0
ISSN 1239-1891