

Paradigmi e tecniche di valutazione dell'usabilità

II parte

VALUTAZIONE DELL'IMPLEMENTAZIONE

Si distingue dalla valutazione del design in quanto:

- Richiede un prototipo del sistema, o una completa implementazione;
- Coinvolge gli utenti del prodotto.
- Metodi empirici: valutazione sperimentale
- Metodi osservazionali: tecniche di interrogazione che richiedono feedback dell'utente

DECIDE:

Un framework per guidare alla valutazione

- Determinare gli *obiettivi* che la valutazione intende perseguire.
- Esplorare le specifiche *domande* cui si deve rispondere.
- Scegliere (Choose) il *paradigma* e le *tecniche di valutazione* per rispondere alle domande.
- Identificare le *questioni pratiche*.
- Decidere come trattare le *questioni etiche*.
- Valutare (Evaluate), interpretare e presentare i *dati*.

Determinare gli obiettivi

- Quali sono gli obiettivi ad alto livello della valutazione?
- Chi la vuole e perchè?
- Gli obiettivi influiscono sul paradigma scelto per lo studio.
- Alcuni esempi di obiettivi:
 - Identificare la migliore metafora su cui basare il design.
 - Controllare che l'interfaccia finale sia coerente.
 - Investigare come la tecnologia influenza le pratiche di lavoro.
 - Migliorare l'usabilità di un prodotto esistente.

Il Caso di studio del SLN4MOP

Goals of the evaluation:

- To evaluate the general reaction of farmers against a mobile application supporting their daily activities
- To perform an evaluation study specifically on the crop selection functionality
- To estimate how the designed interface is perceived by potential users
- Identify possible problems

valutazione dell'implementazione >>

Esplorare le domande

- *Tutte* le valutazioni necessitano di obiettivi e domande che le guidino in modo da non perdere tempo su studi mal impostati.
- Per esempio, l'obiettivo di capire perchè molti clienti preferiscono acquistare biglietti aerei cartacei piuttosto che quelli elettronici, può essere spezzato in sottoproblemi:
 - Che atteggiamento hanno i clienti verso questi nuovi biglietti?
 - Si preoccupano per la sicurezza?
 - L'interfaccia per ottenerli è povera?

Il Caso di studio del SLN4MOP

Obiettivo:

- To estimate how the designed interface is perceived by potential users

Domande

- Are the chosen metaphors appropriate?
- Are the chosen layouts appropriate?
- Is the visualized information appropriate?

valutazione dell'implementazione >>

Scegliere paradigma e tecniche di valutazione

- Quale paradigma?
 - Valutazione 'quick and dirty' (feedback rapido dall'utente su questioni specifiche)
 - Indagine di laboratorio
 - Indagine su campo
 - Valutazione euristica
- Il paradigma di valutazione influenza fortemente le tecniche usate, e come i dati sono analizzati e presentati.

Es. indagini su campo non prevedono testing e modellizzazione.

Il Caso di studio del SLN4MOP

Il metodo di valutazione:

- Observational study- ThinkAloud Cooperativo
- A Follow-up Interview

Field study – Dec 2012

- SLN application was field tested with 32 farmers at two different locations in Sri Lanka.
- Field study was done in the farmer's native language: Sinhala.
- Three parts of the study:
 - Survey to gather the general demographic details of the farmers
 - Use of the prototype
 - Open-ended questionnaire to gather farmer experience of using the prototype, usability issues and other relevant information

Procedure

- Brief Introduction to the Project



11

Il Caso di studio del SLN4MOP

General Instructions for the Farming Prototype

In the next 20 minutes or so, you will be carrying out 3 tasks with the Farming prototype, which are related to the crop selection activity. Each of you will perform different tasks within a common scenario reproducing the situation when a farmer has make a decision on which is the most convenient crop to seed, for a given land and a given season. Note that we intentionally leave out some of the detailed task steps so that we can determine how well the system can guide your interactions with it. If you are confused at any point, please make your best guess on how to proceed, using the information that you have been given. We will intervene if necessary to help you make progress.

At the start of each task, please say loud: "Beginning Task" followed by the number of the task. When you are done, please say: "Task Complete." Also, please remember to *think out loud* as you work. It is very important for us to understand your goals, expectations, and reactions as you work through the task. Any further questions?

- ## Procedure cont.
- Individual Farmer evaluation (1-3 farmers @ time)



13

Procedure cont.

- General questionnaire
- Time to play with the mobile phone
- Crop Planner
 - Introduction
 - Demonstration
 - User tasks
- Profit Calculator
 - Introduction
 - Demonstration
 - User tasks



14

Specific Task Instructions

(note to the evaluator: *The first set of 3 tasks will be performed by the first user of group 1, by the first user of group 2 and so on. The second set by the second user of each group. The third set by the third user of each group, and so on.*)

Task 1.1:

Find out whether it is convenient to plan a tomato production for the next season and if so select the quantity you wish to grow.

Task 1.2:

Find out which product is the most convenient for the next season among P1, P2, P3*. Make your decision about the quantity to grow.

Task 1.3:

Browse the historical data and make your comments loud

...

valutazione dell'implementazione >>

Identificare questioni pratiche

Per esempio, come:

- selezionare gli utenti
- rientrare nel budget
- rispettare i tempi
- trovare i valutatori
- selezionare le attrezzature

Sample Characteristics Demography

Respondents (32) based on					
Region		%	Education		%
Galewela	14	44%	Primary	2	6%
Dambulla	18	56%	Secondary (O/L)	22	69%
Age		%	Secondary (A/L)	5	16%
< 21	0	0%	Diploma	0	0%
21 - 30	4	13%	Graduate	2	6%
31 - 40	9	28%	Post Graduate	0	0%
41 - 50	11	34%	No Proper Education	1	3%
> 51	8	25%	Phone Availability		%
Sex		%	Yes	26	81%
Male	32	100%	No	6	19%
Female	0	0%	2 Phones	9	28%
			3 Phones	1	3%

17

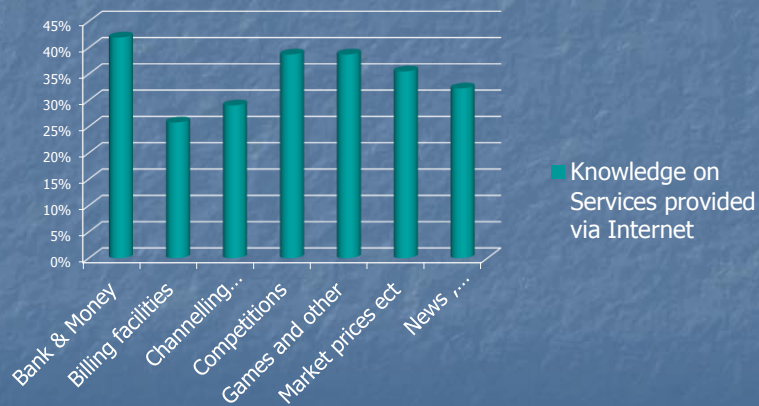
Technology Usage Computer/Internet usage

Computer / Internet Usage		%
Computer Users	8	25%
Non- Computer Users	24	75%
Internet Users	3	9%
Non- Internet Users	29	91%
use internet via other people	2	6%
Not Heard about Internet	1	3%

18

Technology Usage cont...

Knowledge on Services provided via Internet



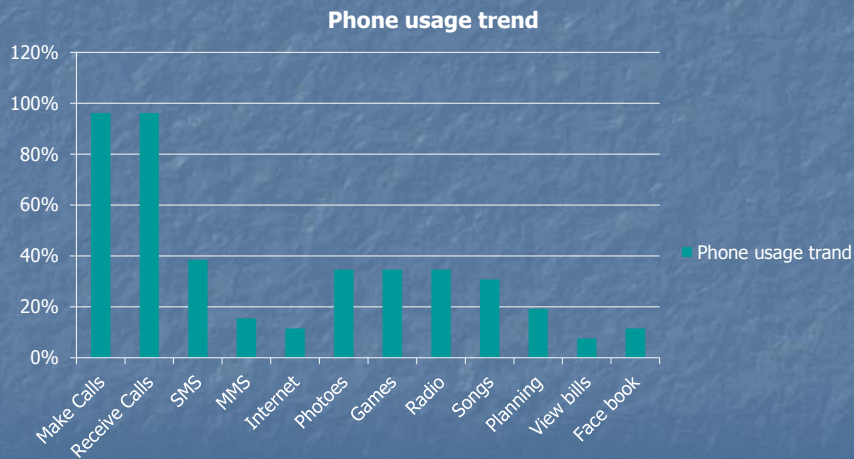
19

Sample Characteristics Demography

Respondents (32) based on					
Region		%	Education		%
Galewela	14	44%	Primary	2	6%
Dambulla	18	56%	Secondary (O/L)	22	69%
Age		%	Secondary (A/L)	5	16%
< 21	0	0%	Diploma	0	0%
21 - 30	4	13%	Graduate	2	6%
31 - 40	9	28%	Post Graduate	0	0%
41 - 50	11	34%	No Proper Education	1	3%
> 51	8	25%	Phone Availability		%
Sex		%	Yes	26	81%
Male	32	100%	No	6	19%
Female	0	0%	2 Phones	9	28%
			3 Phones	1	3%

20

Technology Usage cont...



21

valutazione dell'implementazione >>

Valutare, interpretare & presentare i dati

- Come vengono analizzati e presentati i dati dipende dal paradigma e dalle tecniche usate.
- Bisogna considerare anche quanto segue:
 - Affidabilità: lo studio può essere replicato?
 - Validità: sta misurando ciò che avevate pensato?
 - Alterazioni: il processo sta creando alterazioni?
 - Scalabilità: è possibile generalizzare i risultati?
 - Validità ambientale: l'ambiente in cui si porta avanti lo studio ha influenza su di esso?

Testing the prototype - issues



- How to use a smart phone - Understanding the difference between pressing the keys on a normal mobile phone and touching a screen on a smart phone
- How to use an application on a smart phone
- What to do next in an application
- Selecting items from a drop down box
- Concepts behind login and saving

Crop Planner – Evaluation

- Farmer tasks
 - Select a crop type.
 - Compare two or more crops.
 - Check the history.



Crop Planner – Evaluation

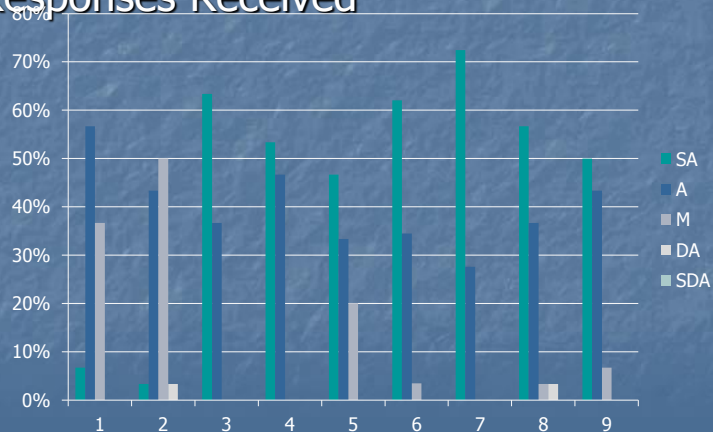
■ Interview Questions (5 likert Scale)

1	All information for the crop choosing stage is provided
2	Information is sufficient for decision making
3	Provide knowledge on different crop varieties
4	Knowledge on history is important
5	Market prices of the previous year are important in deciding a crop
6	Crop comparison facility is essential in deciding a crop
7	Information provided using the Color code is clear
8	Color code usage is important in deciding a crop
9	Functionality provided in this system can be easily learnt by anyone

25

Crop Planner – Evaluation

■ Responses Received



26

Crop Planner – Evaluation

- Interview Questions (Open ended)
 1. Name three (03) main factors that you would consider in prior deciding a crop for cultivation.
 2. What three things did you like most about the Farming prototype?
 3. What three things did you like least about the Farming prototype?
 4. If the Farming prototype was made available to you, would you like to use it or not? Why?
 5. What do you suggest as changes to the design of the Farming prototype?
 6. Do you have any other comments or reactions?

27

Crop Planner – Evaluation

Interview Questions (Open ended)

- *Name three (03) main factors that you would consider in prior deciding a crop for cultivation.*

Revenue	53%
Market Prices	41%
Weather Factors	38%
Crop Variety / Seeds Quality	28%
Seasonal Crop types	28%
Water Availability	22%
Demand in the Market	16%
Current Cultivation Trend	13%
Nursery Preparation	13%
Less Potential to Diseases	9%
Other	25%

28

Crop Planner – Evaluation

Interview Questions (Open ended)

- *What three things did you like most about the Farming prototype?*

Usage of Color code	56%
Seasonal crop varieties for the area/ Information regarding crops	47%
Comparison Facility	34%
History	25%
Current production statistics/Percentages	25%
All functionality	9%
Presentation of clear information	6%
Ability to select crops	3%
Ability to get information easily in less time	3%
Language used(sinhala)	3%

29

Crop Planner – Evaluation

Interview Questions (Open ended)

- *What three things did you like least about the Farming prototype?*

N/A	81%
Difficulty in finding the next action	3%
Less services	6%
Limited number of crops	3%
Unavailability of infor w.r.t. pest and diseases	3%
How to safe guard the cultivation	3%

30

Crop Planner – Evaluation

Interview Questions (Open ended)

- *If the Farming prototype was made available to you, would you like to use it or not? Why?*

Knowledge on Agriculture, Farming activities / technical knowledge
Market prices
Ability to learn unknown facts / new knowledge
Ability to access useful information
Ability to identify current growing trend.
Ability to take decisions on time.
Ability to get information from home.
Easy to use.
Save time to access information and knowledge
Can earn a profit
Flexibility
Easy to learn

31

Crop Planner – Evaluation

Interview Questions (Open ended)

- *What do you suggest as changes to the design of the Farming prototype?*

New Facts

New cultivation techniques used by other provinces & countries
Cultivation increase methods
Crops that have a demand at the market
More information about the history
Predicted cost of a crop to be cultivated
Current supply at the market
Reasons for occurring different Pest & Diseases
Cultivated quantities per season based on provinces
Income prediction
A way to sell the Seeds produced among farmers
Water availability

Required weather / climatic conditions
Pest and diseases / Avoidance / Controlling methods & Characteristics to identify
Information for all stages of cultivation
Required pesticides details
Crop varieties (new/hybrid)
Required fertilizer / how to apply
Daily prices in the market
Better classification for crops vs varieties
Complete Information for each crop
Fungus information
Quality seed information / places to buy
Estimated yield quantity per crop variety
Seed requirement per acre
Weather forecast
Period to grow
Applicable crops(Per season/are/soil)
Nursery Preparation

Known Facts

32

Crop Planner – Evaluation

Interview Questions (Open ended)

- *Do you have any other comments or reactions?*

Internet facility

Up to date information

Accuracy

100% workability

linkage between research institutes and other Stake holders

More information w.r.t. farming life cycle

information accessibility with less cost or free

Training

33

valutazione dell'implementazione >>

METODI OSSERVAZIONALI

Think Aloud (*pensare a voce alta*)

- l'utente è osservato durante il task;
- all'utente è chiesto di descrivere cosa sta facendo e cosa egli pensa stia accadendo.

Vantaggi:

- semplicità;
- può fornire utili intuizioni;
- può mostrare come il sistema è realmente usato.

Svantaggio:

- soggettivo

Valutazione Cooperativa

E' una variazione del think aloud.

- il valutatore e l'utente interagiscono durante la sessione di valutazione;
- l'utente è incoraggiato a criticare attivamente il prodotto e non solo usarlo;
- il valutatore può intervenire durante i momenti più critici dell'interazione per indagare sulle difficoltà dell'utente e verificare proposte alternative;

TECNICHE DI INTERROGAZIONE

Si basano sull'idea che il miglior modo per scoprire come il sistema soddisfa le richieste dell'utente è chiederlo all'utente stesso.

Tali tecniche sono:

- soggettive;
- poco costose.

Due tipi di tecniche di interrogazione:

- interviste
- questionari

INTERVISTE

Generalmente seguono un approccio top-down.

Vantaggi.

- Le domande vengono adattate al contesto;
- I problemi possono essere esaminati a fondo.

Svantaggi.

- molto soggettive;
- impiegano molto tempo.

Interviste

- Non strutturate – non sono guidate da un documento scritto. Ricche ma non replicabili.
- Strutturate – sono rigorosamente scritte, spesso come un questionario. Replicabili ma meno ricche.
- Semi-strutturate – guidate da un documento scritto ma questioni interessanti possono essere esplorate in profondità. Possono fornire un buon compromesso tra ricchezza e replicabilità.

Le basi per intervistare

- Ricordate il framework DECIDE
- Obiettivi e domande guidano tutte le interviste
- Due generi di domande:
 - 'domande chiuse' hanno un formato di risposta predeterminato, es., 'sì' o 'no'
 - 'domande aperte' non hanno un formato di risposta predeterminato
- Le domande chiuse sono più veloci e più facili da analizzare

Cose da evitare quando si preparano le domande per un'intervista

- Domande lunghe
- Frasi composte – dividerle in due
- Usare gergo e linguaggio che l'intervistato non può capire
- Domande che fanno assunzioni, es. "Perché ti piace ...?"
- Influenze non intenzionali, nel modo in cui è posta la domanda

Componenti di un'intervista

- *Introduzione* – presentatevi, spiegate gli obiettivi dell'intervista, rassicurate l'intervistato sulle questioni etiche, chiedete di registrare, presentate un modulo di consenso informato.
- *Riscaldamento* – formulate le prime domande semplici e non minatorie.
- *Corpo principale* – presentate le domande in un ordine logico
- Un periodo di *raffreddamento* – includete poche facili domande per allentare la tensione alla fine
- *Chiusura* – ringraziate l'intervistato, segnalate la fine, es., spegnete il registratore.

Il processo dell'intervista

- Usate framework DECIDE come guida
- Vestitevi in modo simile ai partecipanti
- Controllate prima l'apparecchio di registrazione
- Predisponete un sistema per codificare i nomi dei partecipanti che aiuti a preservare un clima confidenziale.
- Siate gentili
- Chiedete ai partecipanti di riempire un modulo di consenso informato

Sondare e suggerire (*probes and prompts*)

Nelle interviste semi-strutturate è possibile in alcuni casi:

- Sondare – mezzo per ottenere più informazioni. es., 'vorresti aggiungere qualcosa?'
- Suggerire – mezzo per aiutare l'intervistato, es., aiuto nel ricordare un nome e far andare avanti l'intervista
- Ricordate che sondaggi e suggerimenti non dovrebbero comunque creare alterazioni.
- Troppo può incoraggiare i partecipanti a indovinare la risposta.

Interviste di gruppo

- Note anche come 'focus groups'
- Tipicamente 3-10 partecipanti
- Forniscono un variegato range di opinioni
- Devono essere gestite per assicurarsi :
 - che ciascuno contribuisca
 - che la discussione non sia dominata da una persona
 - che l'agenda degli argomenti venga coperta

Analizzare i dati dell'intervista

- Dipende dal tipo di intervista
- Interviste strutturate possono essere analizzate come i questionari
- Interviste non strutturate generano dati come le osservazioni da parte dei partecipanti
- È meglio analizzare il più presto possibile interviste non strutturate, per ricavare argomenti e temi dai dati

valutazione dell'implementazione >>

QUESTIONARI

All'utente vengono date un insieme di domande prefissate.

Vantaggi:

- velocità;
- analisi rigorosa.

Svantaggi:

- poco flessibili ed esplorativi.

Questionari

- Le domande possono essere chiuse o aperte
- Le domande chiuse sono le più facili da analizzare e possono essere poste dal computer
- Possono essere somministrati a popolazioni numerose
- Carta, email e web usati per diffonderli
- Il vantaggio dei questionari elettronici è che i dati vanno in un database che ne rende facile l'analisi
- Il campionamento può essere un problema quando la taglia della popolazione non è nota, come accade online

Stile del questionario

- Varia a seconda dell'obiettivo, perciò è bene usare il framework DECIDE come guida
- I formati di un questionario includono:
 - checkbox 'sì', 'no'
 - checkbox che offrono molte scelte
 - scale Likert di valutazione
 - scale semantiche
 - risposte incomplete
- Le scale Likert hanno range di punti
 - I punti di scala 3, 5, 7 e 9 sono comuni
- Pareri discordanti su quali scale siano migliori,
 - Es. A pochi punti o a molti punti
 - Con numero pari o numero dispari di punti

Sviluppare un questionario

- Fornire una chiara dichiarazione di intenti e garantire l'anonimato dei partecipanti
- Pianificare le domande – se si sta sviluppando un questionario basato sul web, bisogna prima progettare off-line
- Decidere se le frasi saranno tutte positive o tutte negative o un mix
- Testare le domande (*pilot testing*) – sono chiare, c'è spazio sufficiente per le risposte?
- Decidere come verranno analizzati i dati e consultare un esperto di statistica se necessario

User Reactions Survey

Now that you have completed the tasks, we would like to know some of your reactions, both in general and to specific features of the system.

Name: _____

- What three things did you like most about the Farming prototype?
- What three things did you like least about the Farming prototype?
- If the Farming prototype was made available to you, would you use it or not? Why?

Please respond to the following 10 items by circling the opinion that best corresponds to your own

- The list of functionalities provided by the application follows crucial steps of the farming process
- Strongly Disagree Disagree Neutral
 Agree Strongly Agree
- The crop selection functionality allows me to search for crop varieties faster
- The use of the traffic light metaphor (red/yellow/green buttons) to classify the categories of production levels was clear to me
- The purpose of the history button is clear

User Reactions Survey (cont.)

- Gaining information about amounts and prices of certain products in past seasons does help making a decision
- The information supplied during the crop selection process does help making a decision
- The comparative analysis of different products is a useful alternative to the direct selection of the product to seed.
- Once I made my decision and selected a crop quantity, the effects of my decision were visible on the interface.
- It was easy to realize that a product may change the associated color from green to yellow/red and from yellow to red, any time I or other users make a decision on that product.
- Overall, the crop selection functionality enhances the effectiveness of my decision at this stage of the farming process
- What do you suggest as changes to the design of the Farming prototype?
- Do you have any other comments or reactions?

Thank you for your participation!

Incoraggiare una buona risposta

- Assicurarsi che lo scopo dello studio sia chiaro
- Promettere l'anonimato
- Assicurarsi che il questionario sia ben progettato
- Offrire una versione breve per quelli che non hanno il tempo di compilare un questionario lungo
- Se inviato per posta, includere una busta affrancata con l'indirizzo per la restituzione
- Seguire con email, telefonate, lettere
- Fornire un incentivo
- Il 40% di risposte è alto, il 20% è spesso accettabile

Vantaggi dei questionari online

- Spesso le risposte si ricevono velocemente
- Nessun costo di fotocopie e spedizione
- I dati possono essere raccolti in database per l'analisi
- Il tempo richiesto per l'analisi dei dati è ridotto
- Gli errori possono essere corretti facilmente

Problemi con i questionari online

- Campionamento problematico se la taglia della popolazione è ignota
- Evitare che le persone rispondano più di una volta
- Qualche volta è capitato che gli intervistati hanno cambiato le domande in questionari inviati per posta elettronica.

Analisi e presentazione dei dati dei questionari

- Presentare i risultati in modo chiaro – le tabelle possono aiutare
- Semplici statistiche possono dire molto, es. media, mediana, moda, deviazione standard
- Le percentuali di risposte in una data categoria sono utili solo nel caso di molti partecipanti, per standardizzare i dati, es. Per confrontare più insiemi di risposte.
- Per una scarsa popolazione il numero reale di risposte è più onesto.
- Grafici a barre mostrano bene dati categoriali
- Statistiche più avanzate possono essere usate se necessario

■ QUIS (*Questionnaire for User Interaction Satisfaction*)

Es. - Messages which appear on screen:

confusing									clear	
1	2	3	4	5	6	7	8	9	NA	

- Instructions for commands or choice:

confusing									clear	
1	2	3	4	5	6	7	8	9	NA	

■ SUMI (*Software Usability measurement Inventory*)

■ MUMMS (*Measuring the Usability of Multi-Media Systems*)

<http://www.ucc.ie/hfrg/questionnaires/>

SUMI (*Software Usability Measurement Inventory*)

Sviluppato dal collegio universitario di Cork nell'ambito del progetto MUSIC (Measuring the Usability of Systems in Context/Metrics for Usability Standards in Computing).

E' un questionario internazionale standardizzato di 50 item, disponibile in 5 lingue (Inglese, Tedesco, Olandese, Spagnolo e Italiano).

Richiede approssimativamente 5-10 minuti per essere completato, e almeno 10 soggetti rappresentativi per fornire risultati affidabili.

Le misure sviluppate da SUMI sono basate principalmente sui seguenti aspetti:

- ✓ Efficienza;
- ✓ Conoscenza.

SUMI (*Software Usability Measurement Inventory*)

1. Questo software risponde troppo lentamente agli input?
2. Raccomanderei questo software ai miei colleghi?
3. Le istruzioni e i prompt sono di aiuto?
4. Il software si è qualche volta fermato inaspettatamente?
5. ...

PROTOCOLLI DI ANALISI

Per registrare le azioni dell'utente possono essere utilizzati i seguenti metodi:

- Foglio e penna;
- Registratori audio;
- Registratori Video;
- Computer logging;
- Appunti dell'utente;

Il Caso di studio del SLN4MOP

Data Collection Form for SLN4MOP prototype

Date: _____ Participant ID: _____
Evaluator: _____

Task number: _____ Start time _____
Stop time _____

Comments made by participant:

Errors or problems observed (including assistance offered):

Other relevant observations:

TOOL AUTOMATICI PER L'ANALISI DEI PROTOCOLLI.

DRUM (*Diagnostic Recorder for Usability Measurement*)

E' uno strumento software sviluppato dalla NPL (National Physical Laboratory – Middlesex, UK) all'interno del progetto MUSIC.

Serve per velocizzare l'analisi dei filmati registrati sull'interazione dell'utente con il software e per aiutare a gestire la valutazione dell'usabilità nel suo complesso.

I valutatori scelgono gli eventi che vogliono controllare, e tali eventi vengono registrati automaticamente sul nastro e visualizzati sul video per essere analizzati.

valutazione dell'implementazione >>

Metodi empirici: Valutazione sperimentale

- sono effettuate valutazioni controllate di specifici aspetti del comportamento interattivo.
- i valutatori scelgono le ipotesi da testate;
- un numero di condizioni sperimentali sono considerate, le quali differiscono solo nel valore di alcune variabili controllate;

Fattori Sperimentali

Soggetti scelti.

La scelta dei soggetti è vitale nel successo dell'esperimento.

Essi devono essere scelti tenendo in considerazione l'utente finale del prodotto.

Il numero non deve essere inferiore a 10.

Variabili da testare.

- variabili indipendenti (IV) – caratteristiche cambiate per produrre differenti condizioni;
- variabili dipendenti (DV) – caratteristiche misurate nell'esperimento. Es. tempo trascorso, numero di errori.

Ipotesi

- E' una predizione dei risultati degli esperimenti.
- E' formulata in termini di variabili indipendenti e dipendenti, affermando che una variazione nella variabile indipendente causerà una variazione nella variabile dipendente.
- Lo scopo dell'esperimento è mostrare che la predizione è corretta. Ciò è fatto invalidando l'**ipotesi nulla**, la quale stabilisce che le variabili dipendenti non variano in funzione delle variabili indipendenti.

Design sperimentale

➤ *within groups design*: ogni soggetto esegue gli esperimenti sotto tutte le condizioni. L'esperimento può soffrire di trasferimenti di conoscenze, ma questi possono essere diminuiti se l'ordine nel quale le condizioni sono affrontate è variato tra gli utenti. (n.b. errata corregge testo di Dix pag. 298 "Il secondo disegno... sotto tutte le condizioni")

➤ *between groups design*: ogni soggetto esegue gli esperimenti soltanto sotto una condizione. L'esperimento non soffre di trasferimenti di conoscenze, ma è richiesto un grande numero di utenti.

Analisi dei risultati

I risultati vengono analizzati attraverso dei metodi statistici.

Metodi empirici: Valutazione sperimentale

Cosa può variare (e quali scelte effettuare):

- Individui/gruppi (non solo studenti UNISA!)
- Task/attività
- Prodotti/sistemi
- Principi/teorie
- Conoscenza pregressa ed esperienza
- Effetti dell'apprendimento e dell'ordinamento dei task

Su che cosa si vogliono ottenere informazioni?

Che cosa costituisce invece *rumore*?

una breve storia...

Un lavoro presentato a una importante conferenza internazionale dell'ACM

Un buon articolo empirico

La ricerca di un supporto collaborativo per un task X

Tre moduli software a confronto:

A. Un software sincrono 'domain-specific'

(un software collaborativo per il software project management)

B. Un software sincrono 'generico'

C. Un software asincrono 'generico'

	syncr.	asyncr.
domain spec.	A	?
generic	B	C

L'esperimento

Un numero congruo di partecipanti presi sotto ciascuna condizione

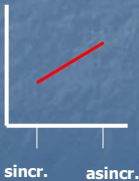
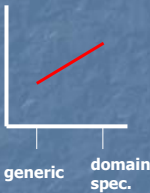
Metriche di qualità

Risultati significativi con $p < 0,05$

domain spec. > generic

asincrono > sincrono

	syncr.	asyncr.
domain spec.	A	★
generic	B	C



Cosa c'era di sbagliato?

Gli effetti dell'interazione

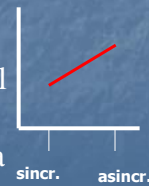
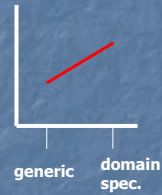
- Il gap è interessante da studiare ma ...
- ... non sempre risulta essere la migliore soluzione

NON si sono usate variabili indipendenti

- 3 software completamente diversi
- come se fosse un esperimento su 3 persone!

supponete ad esempio che il software B sia stato mal progettato. In quel caso non sarebbe la caratteristica di essere generico la causa di una cattiva reazione da parte degli utenti.

	syncr.	asincr.
domain spec.	A	?
generic	B	C



UN ESEMPIO DI VALUTAZIONE SPERIMENTALE

CONCLUSIONI

Nella scelta del metodo di valutazione alcuni fattori da considerare possono essere:

- quando è effettuata la valutazione? (design o implementazione);
- Quale stile di valutazione è richiesto? (laboratorio, campo);
- La tecnica quanto deve essere obiettiva?
- Quante risorse sono disponibili? (tempo, soggetti, esperti)

Studi pilota

- Una piccola prova dello studio che si intende portare avanti, che è opportuno fare prima di addentrarsi nello studio vero e proprio.
- Lo scopo è di assicurarsi che il proprio piano di valutazione sia ammissibile.
- Gli studi pilota controllano:
 - che chi esegue lo studio sia in grado di condurre il processo di valutazione
 - che i testi delle interviste, i questionari, gli esperimenti ecc., funzionino appropriatamente
- Val la pena farne diversi per mettere a punto problemi prima di eseguire lo studio vero e proprio.
- Chiedete ai colleghi se non riuscite a dedicare allo studio pilota utenti reali.

Punti chiave della valutazione dell'implementazione

- Interviste strutturate, non strutturate, semi-strutturate, focus groups, questionari
 - Domande chiuse sono più facili da analizzare e si possono replicare
 - Domande aperte sono più ricche
 - Check boxes, scale Likert e scale semantiche
- Valutazione da parte di esperti: euristiche e walkthrough
 - Relativamente economiche perchè nessun utente è coinvolto
 - Valutazione euristica relativamente facile da apprendere
 - Può trascurare problemi chiave e identificare falsi problemi

Per casa ...

- Attività: "acquistare due biglietti per il cinema su <http://www.thespacecinema.it/>"
- Applicate il cognitive walkthrough per la valutazione di quest'attività

Riferimenti

- Alan Dix, Janet Finlay, Gregory Abowd, Russell Beale, "Interazione Uomo Macchina", Cap. 8