



UNIVERSITÀ DEGLI STUDI DI SALERNO
DIPARTIMENTO DI INFORMATICA

LEZIONE

Paradigmi e tecniche di valutazione dell'usabilità– Parte 2

Anno Accademico 2022/2023

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.

Quali aspetti dell'usabilità considerare?



(a) Fitbit Charge e (b) Polar A370

- Quanto comodo da indossare?
- Quanto chiara è l'informazione presentata?
- Quali altri info sono presentate?
- Quanto dura la batteria?
- **Quanto accurato è il dispositivo?**
- Quanto alla moda?

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 anziani 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?

C- SCEGLIERE PARADIMA 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.

IDENTIFICARE QUESTIONI PRATICHE

Per esempio, come:

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

DECIDERE SU QUESTIONI ETICHE

- Sviluppare un modulo di consenso informato
- I partecipanti hanno il diritto di:
 - conoscere gli obiettivi dello studio
 - cosa succederà ai risultati dello studio
 - privacy dei dati personali
 - non essere citati senza il proprio consenso
 - abbandonare lo studio in qualsiasi momento
 - essere trattati educatamente

E- 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?

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.

Es. di ipotesi nulla Il numero medio di errori commessi eseguendo il task X con il prototipo A è uguale al numero medio di errori commessi eseguendo lo stesso task col prototipo B.

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.

Esempio

Progettare un esperimento per testare se colorare gli elementi del menu durante la selezione aumenta la precisione.

Soggetti: Scelti tra gli utenti finali del prodotto;

Ipotesi: Colorare le selezioni rende la scelta più accurata;

(ipotesi nulla: colorare le selezioni non fa aumentare il livello di accuratezza nella scelta)

Variabile Indipendente: Colore;

Variabile Dipendente: Numero di errori;

Design Sperimentale: Between groups.

Task: Le interfacce sono identiche tranne per la colorazione degli item presente solo nella seconda. Ai soggetti è presentata una schermata di un menu e a voce sono tenuti a dire quale scelta hanno selezionato. Ogni presentazione pone gli elementi in posizioni diverse.

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 (cont.)

- **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., interrompete la registrazione.

IL PROCESSO DELL'INTERVISTA

- Usate framework DECIDE come guida
- Vestitevi in modo simile ai partecipanti
- Controllate prima lo strumento 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

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 'si', '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 progettarlo 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

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
- 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

ALCUNI QUESTIONARI PER LA VALUTAZIONE DELL'USABILITA'

- **QUIS** (*Questionnaire for User Interaction Satisfaction*)
 - esempio
- **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 20 lingue (Inglese, Tedesco, Olandese, Spagnolo, Italiano, Giapponese, Cinese, Norvegese ecc.).

Richiede approssimativamente 5-10 minuti per essere completato, e circa 20 soggetti rappresentativi (non meno di 12) 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. C'è abbastanza informazione sull'interfaccia quando ce ne è bisogno?
6. Quando uso questo software, penso di avere tutto sotto il mio controllo?
7. Questo software funziona bene quando voglio fare cose non standard?
8. ...

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;

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.

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

- Coinvolge gli utenti finali del prodotto
- Tecniche sperimentali e/o osservazionali e interviste/questionari
- Interviste strutturate, non strutturate, semi-strutturate, focus groups, questionnaire
 - ❖ 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 problem chiave e identificare falsi problemi

Esercitiamoci

- - Attività:
“acquistare due biglietti per Berlino <https://www.flixbus.it/>
- - Spiegate come impostereste uno studio sperimentale utilizzando il framework DECIDE.

Riferimenti

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