

Paradigmi e tecniche di valutazione dell'usabilità

OBIETTIVI DELLA VALUTAZIONE

- Stimare le funzionalità offerte dal sistema;
- Stimare gli effetti dell'interfaccia sull'utente;
- Identificare specifici problemi.

Due stili di valutazione principali

- Indagini di Laboratorio
- Indagini sul campo

stili di valutazione >>

INDAGINI DI LABORATORIO

Il sistema viene studiato all'interno di un laboratorio.

Solitamente coinvolge soltanto il designer e i valutatori anche se gli utenti possono essere portati all'interno del laboratorio per prendere parte agli studi di valutazione.

Vantaggi:

- disporre di attrezzatura specifica (telecamere, registratori audio, meccanismi di logging etc.);
- ambiente libero da interruzioni.

Svantaggi:

- perdita del contesto;
- difficoltà nell'osservare la cooperazione tra diversi utenti;

Appropriato:

- se la posizione del sistema è pericolosa o impraticabile;
- per task per singolo utente, che presentano molti vincoli;
- per manipolare deliberatamente il contesto per scoprire problemi o osservare procedure meno usate.
- per confrontare design alternativi in un contesto controllato

INDAGINI SUL CAMPO

La valutazione viene effettuata nell'ambiente di lavoro dell'utente per osservare il sistema in azione.

Vantaggi:

- ambiente naturale;
- mantenimento del contesto (anche se le osservazioni possono alterarlo);
- possibilità di studi che durano nel tempo;

Svantaggio:

- distrazioni (alto livello del rumore, suono del telefono, etc.);

Appropriato:

- quando il contesto è cruciale;
- per osservazioni a lungo termine;

VALUTAZIONE DEL DESIGN

- Viene effettuata in fase di design prima che sia iniziata la fase di implementazione. In questo modo possono essere evitati grandi sprechi di risorse.
- Solitamente i metodi di valutazione del design non coinvolgono l'utente ma vengono effettuati dal designer e da esperti valutatori.
- Possono essere applicati anche nel processo di sviluppo su prototipi o su una versione completa del sistema.

Chiedere agli utenti e agli esperti

- 1. Interviste e questionari**
- 2. Valutazione del design**

Chiedere agli esperti

- Gli esperti usano la loro conoscenza degli utenti e della tecnologia per rivedere l'usabilità del software
- Le critiche degli esperti possono essere rapporti formali o informali
- I walkthrough richiedono di attraversare uno scenario pre-pianificato notando potenziali problemi
- La valutazione euristica è una revisione guidata da un insieme di euristiche

Quattro possibili approcci

- Cognitive Walkthrough
- Valutazione euristica
- Valutazione basata su revisioni
- Valutazione basata sull'uso di modelli

Cognitive Walkthrough (*Sondaggio Cognitivo*)

Basato sulla teoria dell'apprendimento esplorativo di Polson.

- Valuta quanto il design supporta l'utente nell'apprendimento dei task;
- usualmente è effettuato da esperti in psicologia cognitiva;
- esperti usando principi psicologici analizzano il design per identificare potenziali problemi.
- deriva dal 'code walkthrough', tecnica usata in ingegneria del software, riferito a una sequenza di passi che rappresentano un segmento di codice di programma

Per effettuare un walkthrough cognitivo bisogna avere:

1. Descrizione del prototipo del sistema;
2. Descrizione di un task rappresentativo che l'utente effettua sul sistema;
3. Lista completa delle azioni necessarie per completare il task utilizzando il prototipo;
4. Indicazioni di chi sono gli utenti del prodotto ed il loro tipo di esperienza e di conoscenza;

➤ I valutatori seguono passo per passo la sequenza di azioni descritta al punto 3

➤ Per ogni azione dovranno rispondere alle seguenti domande:

- ❖ L'utente tenterà di produrre l'effetto che produce l'azione?
- ❖ L'utente noterà che è disponibile sull'interfaccia la corretta azione da eseguire per raggiungere l'obiettivo del compito?
- ❖ Se l'utente troverà l'azione corretta sull'interfaccia, saprà che è quella giusta per ottenere l'effetto che sta tentando produrre?
- ❖ Una volta eseguita l'azione, l'utente comprenderà il feedback che ottiene? Assocerà il risultato dell'azione con il conseguimento dell'obiettivo?

Un Esempio: valutazione del design di un videoregistratore

Gli esperti cominceranno col preparare la documentazione necessaria specificando:

1. La descrizione del prototipo del sistema videoregistratore;
2. Descrizione di un task rappresentativo che l'utente effettua sul sistema:

Programmare tramite telecomando un videoregistratore per registrare un programma sul canale 4 che inizia alle 18:00 e finisce alle 19:15 e viene trasmesso il 28 novembre 2013

3. Lista completa delle azioni necessarie per completare il task utilizzando il prototipo;
4. Indicazioni di chi sono gli utenti del prodotto ed il loro tipo di esperienza e di conoscenza;

Un Esempio (cont.) valutazione del design >>

Task: *Programmare tramite telecomando un videoregistratore per registrare un programma sul canale 4 che inizia alle 18:00 e finisce alle 19:15 e viene trasmesso il 28 novembre 2013.*

Azione A: Premere il bottone 'registrazione programmata'.

Risposta A: Il display visualizza la modalità timer.

Azione B: Premere i numeri 1800.

Risposta B: Ogni numero è visualizzato quando è digitato, ed il cursore lampeggiando si muove nella successiva posizione.

Azione C: Premi il pulsante 'registrazione programmata'

Risposta C: Il cursore lampeggiante si muove su 'Fine.'

Azione D: Premere i numeri 1915.

Risposta D: Ogni numero è visualizzato quando è digitato, ed il cursore lampeggiando si muove nella successiva posizione.

Azione I: Premere il tasto ‘registrazione programmata’.

Risposta I: Il numero di programmazione nell'angolo in alto a destra del display inizia a lampeggiare.

Azione J: Premere il tasto ‘trasmissione’.

Risposta J: I dettagli sono trasmessi al videoregistratore ed il display ritorna alla modalità normale.

Esempio.

Orario:

21:45


Canale:

3

1

2

3



4

5


6


7


8


9


0














(a)








Inizio:

1

Fine:

Canale:

Data:

1	2	3	
4	5	6	
7	8	9	0
			
			

(b)

Azione A: Premere il bottone 'registrazione programmata'.

Domanda 1: L'utente tenterà di produrre l'effetto che produce l'azione?

L'interfaccia non dà indicazioni sul fatto che l'utente ha bisogno di premere 'registrazione programmata'. Comunque è plausibile pensare che un utente che abbia dimestichezza col videoregistratore sappia che ciò è richiesto

Domanda 2: L'utente noterà che è disponibile sull'interfaccia la corretta azione da eseguire per raggiungere l'obiettivo del compito?

Il pulsante 'registrazione programmata' è visibile sul telecomando

Azione A: Premere il bottone 'registrazione programmata'.

gn >>

Domanda 3: Se l'utente troverà l'azione corretta sull'interfaccia, saprà che è quella giusta per ottenere l'effetto che sta tentando produrre?

Non è chiaro quale pulsante è quello per la registrazione programmata. L'icona di un orologio (quarto pulsante a destra) è un possibile candidato, ma potrebbe anche essere un pulsante per modificare l'ora. Altri possibili candidati possono essere il quarto pulsante a sinistra o quello con un cerchio pieno (associato alla registrazione). Di fatto l'orologio è la scelta giusta, ma a questo punto è possibile che l'utente fallisca.



Domanda 4: Una volta eseguita l'azione, l'utente comprenderà il feedback che ottiene? Assocerà il risultato dell'azione con il conseguimento dell'obiettivo?

Una volta premuto il pulsante il display va nella modalità 'registrazione programmata' e mostra voci familiari (inizio, fine, canale, data). È ragionevole assumere che l'utente le riconoscerà come segnali di un corretto completamento della prima azione

Un Esempio (cont.)

Gli esperti terranno un foglio a parte dove annotare i problemi che via via vengono osservati.

L'elenco di tali problemi verrà alla fine consegnato ai progettisti che ne esamineranno l'importanza e decideranno se e come intervenire per migliorare il design.

Walkthrough pluralistico

- Variazione del cognitive walkthrough
- Eseguito da un team attentamente gestito
- Il collegio di esperti inizia a lavorare separatamente
- Segue una discussione organizzata che porta a una decisione collegiale
- L'approccio si presta bene al participatory design

VALUTAZIONE EURISTICA

Proposto da Nielsen e Molich (1990).

Ha lo scopo di verificare se l'interfaccia del prodotto rispetta i principi fondamentali dell'usabilità.

Si tratta di un metodo ispettivo che non prevede il coinvolgimento degli utenti finali del prodotto, ma si basa sul giudizio di più esperti di usabilità, che fanno una critica dell'interfaccia, per individuarne eventuali problemi.

Per effettuare una valutazione euristica:

- sono identificati i criteri di usabilità, basati su linee guida e principi di progettazione dell'interfaccia;
- Il design è esaminato da esperti per vedere se i criteri sono violati;

Da studi effettuati da Nielsen, si rileva che 5 esperti diagnosticano circa il 75% dei problemi.

Le euristiche di Nielsen

1. Visibilità dello stato del sistema
2. Corrispondenza tra sistema e mondo reale
3. Controllo e libertà dell'utente
4. Consistenza e standard
5. Prevenzione degli errori
6. Riconoscimento piuttosto che ricordo
7. Flessibilità ed efficienza di utilizzo
8. Design estetico e minimalista
9. Aiutare gli utenti a riconoscere, diagnosticare e recuperare dagli errori
10. Aiuto e documentazione

3 fasi per effettuare la valutazione euristica

- Sessione informativa per dire agli esperti cosa fare
- Periodo di valutazione di 1-2 ore in cui:
 - ogni esperto lavora separatamente
 - un primo passaggio per avere un'idea del prodotto
 - un secondo passaggio per concentrarsi su aspetti specifici
- Sessione in cui gli esperti lavorano insieme per dare priorità ai problemi

Vantaggi e problemi

- Poche questioni etiche e pratiche da considerare
- Può essere difficile e costoso trovare gli esperti
- Gli esperti migliori hanno conoscenza del dominio di applicazione e degli utenti
- Maggiori problemi
 - problemi importanti possono essere trascurati
 - vengono spesso identificati molti problemi banali

VALUTAZIONE BASATA SU REVISIONI

- Basare la valutazione sul risultato di esperimenti analoghi già documentati e presenti in letteratura.
 - ✦ In particolare per questioni generali che si applicano a una varietà di situazioni (es. l'usabilità di diversi tipi di menù, il richiamo di nomi di comandi e la scelta di icone)
- Tuttavia è necessario tener presente che i risultati sperimentali non valgono in qualsiasi contesto
- La revisione deve tener conto delle similitudini e delle differenze tra il contesto dell'esperimento e il design che si sta considerando per la valutazione.

VALUTAZIONE BASATA SULL'USO DI MODELLI

- Alcuni modelli cognitivi e di design forniscono un mezzo per combinare specifiche di progetto e valutazione nello stesso framework
 - ❖ Il modello GOMS (Goals Operators Methods and Selection), che predice le prestazioni dell'utente con una particolare interfaccia e può essere usato per filtrare delle particolari scelte di progetto
 - ❖ Il modello 'keystroke-level", che fornisce previsioni del tempo che l'utente impiegherà a eseguire task fisici di basso livello

VALUTAZIONE DELL'IMPLEMENTAZIONE

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

valutazione dell'implementazione >>

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.

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?

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.

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

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 testare;
- 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.

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;

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