

System Design

Carmine D'Angelo

A.A. 2022/2023

Indice

| | |
|--|-----------|
| 1 Human Computer Interaction (HCI) | 4 |
| 1.1 Sketchpad | 4 |
| 1.1.1 Esempio di un esperimento di una procedura di design | 4 |
| 1.1.2 Nascita dell'HCI e ACM SIGCHI | 5 |
| 2 The Human Factor | 6 |
| 2.1 Sensi Umani | 6 |
| 2.1.1 Occhio (vista o sight) | 6 |
| 2.1.2 Udito (audition) | 7 |
| 2.1.3 Tutto | 7 |
| 2.1.4 Gusto (taste) e olfatto (smell) | 7 |
| 2.2 Responders | 8 |
| 2.2.1 Handedness | 8 |
| 2.2.2 Voce umana | 8 |
| 2.2.3 NVVI | 9 |
| 2.2.4 Occhio usato come responder | 9 |
| 2.2.5 Il cervello | 9 |
| 2.3 Human Performance | 11 |
| 2.3.1 Speed-accuracy Trade-off | 11 |
| 2.3.2 Human Diversity | 11 |
| 2.3.3 Reaction Time | 11 |
| 2.3.4 Skilled Behaviour | 11 |
| 3 Interaction Elements | 12 |
| 3.1 Hard e Soft controls | 12 |
| 3.2 Relazioni tra control e display | 12 |
| 3.3 Mapping Spaziale | 12 |
| 3.3.1 Axis Labeling | 13 |
| 3.3.2 Third Tier | 13 |
| 3.3.3 Trasformazione spaziale: 3D | 13 |
| 3.3.4 Trasformazione spaziale: CD Gain (CD-Ratio) | 14 |
| 3.4 Propeery sensed e control | 14 |
| 3.5 Metafora dell'orologio | 14 |
| 3.6 Modalità (Mode) | 15 |
| 3.6.1 Mode switching | 15 |
| 3.6.2 Mode visibility | 15 |
| 3.7 Mobile context | 15 |
| 3.7.1 Touch screen | 15 |
| 3.8 Accelerometro | 15 |
| 4 Fondazioni scientifiche | 16 |
| 4.1 Sperimentazione | 16 |
| 4.2 Pubblicazione | 16 |
| 4.3 H-index | 16 |
| 4.4 Digital library | 16 |
| 4.5 Valutazioni giornali | 17 |
| 4.6 Riproducibilità di una ricerca | 17 |
| 4.7 Ricerca vs Ingegneria vs Design | 18 |
| 4.8 Empirical Research | 18 |
| 4.8.1 Observational empirical method | 18 |
| 4.8.2 Experimental Method | 19 |
| 4.8.3 Correlational Method | 19 |
| 4.9 Osservare e misurare | 19 |
| 4.9.1 Nominal Data | 19 |
| 4.9.2 Ordinal Data | 19 |
| 4.9.3 Radio Data | 19 |
| 4.9.4 Interval Data | 20 |
| 4.9.5 Quale funzioni calcolabili si possono usare sulle variabili? | 20 |
| 4.9.6 Internal or External Validity | 20 |

| | |
|--|-----------|
| 5 Design HCI experiment | 21 |
| 5.1 Ethics Approval | 21 |
| 5.1.1 Variabile indipendente | 21 |
| 5.1.2 Variabile dipendente (DP) | 23 |
| 5.1.3 Control variable | 23 |
| 5.1.4 Random variable | 23 |
| 5.1.5 Confounding variable | 24 |
| 5.2 Order effect, counterbalancing | 24 |
| 5.2.1 Latin Squares | 24 |
| 5.2.2 Balanced Latin Squares | 25 |
| 5.2.3 Test conditions with multiple factors | 25 |
| 5.2.4 Other techniques | 26 |
| 5.3 Experiment Task | 26 |
| 5.3.1 Knowledge-based Task | 26 |
| 5.4 Questionari | 27 |
| 5.5 System Usability Scale | 27 |
| 5.6 Longitudinal Studies | 27 |
| 5.7 Esperimento fatto in classe (Tapping VS Gesture Typing on Smartphones) | 28 |
| 5.7.1 Come calcolare la performance delle metriche nella text entry? | 28 |
| 6 Hypothesis testing | 29 |
| 6.1 Statistical procedures | 29 |
| 6.2 Analisi della varianza (ANOVA) | 30 |
| 6.2.1 Come si fa il resoconto (report) della F-statistic | 31 |
| 6.3 Più di due test condition | 32 |
| 6.3.1 Between-subject Designs | 32 |
| 6.4 Post Hoc Comparisons Tests | 33 |
| 6.4.1 Scheffé Post Hoc comparisons | 33 |
| 6.5 Two-way ANOVA | 33 |
| 7 Modeling Interaction | 34 |
| 7.1 Modelli predittivi | 34 |
| 7.1.1 Linear prediction equation | 34 |
| 7.1.2 Fitt's law | 35 |
| 7.1.3 Choice reaction time | 36 |
| 7.1.4 Keystroke-level model (KLM) | 36 |
| 8 Research paper | 39 |

Human Computer Interaction (HCI)

Introduzione Nasce come scienza nel 1980. Nasce in questi anni per via dello sviluppo dei personal computer e quindi si voleva migliorare l'usabilità per gli utenti di tutti i giorni.

Mira a creare sistemi che siano efficienti, sicuri, confortevole ed a volte anche a renderle divertenti per chi le utilizza.

Linea degli eventi fino alla HCI ENIAC è il primo calcolatore e nasce nel 1940. Dal 1920 al 1950 si sono utilizzate le schede perforabili, nel 1976 sono state sostituite dall'editor di testo VI.

ACM ACM è una collezione online di documenti scientifici, libri, conferenze e atti di workshop nel campo dell'informatica e dell'ingegneria informatica.

1.1 Sketchpad

È stato ideato nel 1962 da Sutherland. Si basa sulla tecnica del **direct manipulation** (Termine coniato da Ben Shneiderman). Le caratteristiche principali di sketchpad sono:

1. **Azioni incrementali con rapido feedback**: consente all'utente di apportare modifiche incrementali ai disegni, visualizzando in tempo reale il risultato delle modifiche.
2. **Reversibilità delle azioni**: permette all'utente di annullare e ripetere le azioni compiute in qualsiasi momento, consentendo di correggere gli eventuali errori.
3. **Esporazione**: consente all'utente di esplorare le diverse opzioni di design e di testare le idee, permettendo una maggiore flessibilità nel processo di progettazione.
4. **Correzione sintattica delle azioni**: riconosce automaticamente le azioni dell'utente e fornisce un feedback immediato in caso di errori di sintassi, come ad esempio la creazione di una forma non valida.
5. **Rimpiazzamento del linguaggio con le azioni**: ha introdotto il concetto di manipolazione diretta degli oggetti, permettendo all'utente di manipolare gli elementi del disegno attraverso azioni fisiche come cliccare, trascinare e rilasciare. Questa funzionalità ha eliminato la necessità di scrivere codice o linguaggio di programmazione per creare disegni e forme geometriche.

Direct manipulation : La Direct Manipulation è un paradigma di interazione uomo-computer che permette agli utenti di manipolare direttamente gli oggetti sullo schermo attraverso azioni fisiche, come cliccare, trascinare e rilasciare.

Invenzione del mouse : Il mouse fu inventato da Doug Engelbart nel 1963. Non ci un cazzo perché il brevetto scadde prima della commercializzazione del primo mouse intorno agli anni 80. Ha vinto il premio Turing nel 1997.

Variabile Dipendente In statistica, una variabile dipendente è una variabile che viene misurata o osservata durante un esperimento o un'analisi statistica. Il valore di questa variabile dipende dalle altre variabili coinvolte nello studio, in particolare dalla variabile indipendente.

Variabile indipendente In statistica, una variabile indipendente è un tipo di variabile che viene manipolata o controllata durante un esperimento o un'analisi statistica. È chiamata "indipendente" perché il suo valore non dipende dalle altre variabili coinvolte nello studio.

1.1.1 Esempio di un esperimento di una procedura di design

In questo esperimento abbiamo avuti 13 partecipanti.

- Variabili indipendenti: sono utilizzati 6 metodi di input: mouse, light pen, grafacon, joystick (position-control), joystick (rate-control), knee-controlled lever;
- Variabili dipendenti: tempo di completamento dei task (tempo di accesso + tempo di movimento), frequenza degli errori.
- L'esperimento è svolto in modo Within-subjects e counterbalanced
 - within-subjects indica che tutti i partecipanti coinvolti nello studio vengono sottoposti a tutti gli esperimenti.
 - counterbalanced si riferisce all'ordine in cui sono presentati i progetti, in questo caso l'ordine di presentazione delle diverse condizioni viene bilanciato in modo da minimizzare l'effetto dell'ordine di presentazione sulle misure effettuate.
- Task: premere la spacebar, acquisire il device, posizionare il cursore sul targer, selezionare il target.

Come risultati si è ottenuto che il knee control è il più veloce, ma il mouse è più accurato.

Xerox Star Xerox Star è stato il primo computer commerciale ad aver utilizzato un'interfaccia grafica utente (GUI) basata sul concetto di windows, icone, menù e dispositivi di puntamento (WIMP). Supportava la direct manipulation e il concetto what-you-see-is-what-you-get(WYSIWYG).

L'interfaccia utente era basata su un desktop virtuale, sul quale erano presenti icone che rappresentavano documenti, cartelle e programmi, e finestre che permettevano di visualizzare e manipolare i contenuti.

Xerox Star utilizzava la metafora del desktop (Desktop Metaphor), che consiste nell'adottare concetti derivati dall'ambiente dell'ufficio e trasporli sulla schermata del computer. In pratica, l'interfaccia utente mostrava rappresentazioni grafiche (icone) di oggetti comuni come documenti, cartelle, vassoi e accessori, in modo da facilitare la comprensione e l'uso dell'interfaccia da parte degli utenti. Le metafore (metaphores) sono importanti in HCI (Human-Computer Interaction), poiché gli utenti possono utilizzare le loro conoscenze pregresse di un determinato dominio per comprendere e utilizzare l'interfaccia utente del computer. Inoltre, per aumentare l'usabilità, Xerox Star nascondeva i dettagli tecnici del sistema all'utente, ad esempio aprendo un documento anziché invocare un editor di testo. Questo permetteva agli utenti di concentrarsi sulle attività che volevano svolgere, senza preoccuparsi della complessità tecnica del sistema sottostante.

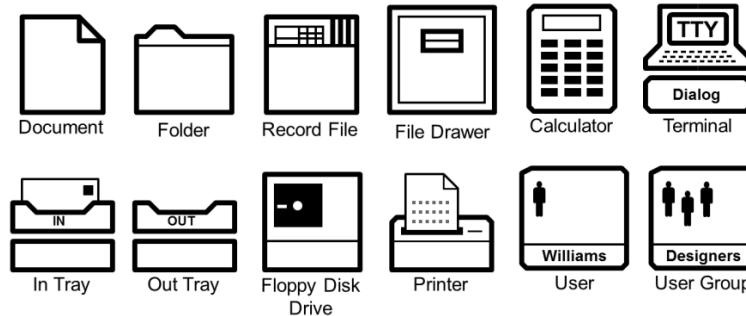


Figura 1.1: Icone GUI Xeros Star

1.1.2 Nascita dell'HCI e ACM SIGCHI

Nel 1983 abbiamo la prima conferenza dell'ACM SIGCHI, nello stesso anno abbiamo la pubblicazione di "The Psychology of Human-Computer Interaction". Nel 1984 Apple annuncia Macintosh. L'HCI diventa importante perché i computer stanno iniziando a diventare di uso comune.

ACM è la più grande associazione di professionisti che lavorano nella ricerca e nella pratica dell'interazione uomo-computer. È composto da molteplici figure (graphic designer, interact designer, antropologi, ecc.) con lo scopo comune di progettare tecnologie utili e usabili seguendo un processo interdisciplinare. Credono che, se fatto correttamente, possa avere il potere di trasformare la vita delle persone.

The Human Factor

Il modello human factor è composto da 3 sezioni:

- **Umano:** il cervello è collegato ai sensori (dispositivi hardware o software utilizzati per rilevare e acquisire dati) e ai responder (componente software che fornisce una risposta a una richiesta da parte dell'utente o dell'interfaccia utente);
- **Interfacce:** composte da display e sensori;
- **Computer:** composti da macchine a stati, display e controlli.

La parte di cui si occupa HCI è quella dei computer. Inizialmente i controlli erano meccanici, negli ultimi anni sono mappati come software mostrati su di un display.

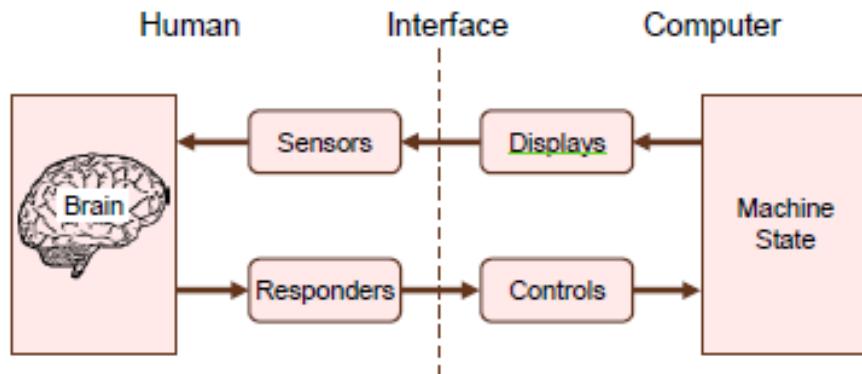


Figura 2.1: Modello human factor

2.1 Sensi Umani

2.1.1 Occhio (vista o sight)

Le persone ottengono l'80% delle loro informazioni dalla vista tramite l'uso dell'occhio. L'azione della vista inizia con la ricezione della luce attraverso la lente (lens). La lente focalizza la luce in un'immagine proiettata sulla retina. La retina è un trasduttore che converte la luce visibile in segnali neurologici inviati al cervello attraverso il nervo ottico.

Vicino al centro della retina si trova la fovea, responsabile della visione centrale nitida, come la lettura o la visione della televisione. L'immagine della fovea nell'ambiente comprende un angolo visivo, approssimativamente equivalente alla larghezza del pollice a distanza di un braccio. Sebbene la fovea sia grande solo l'1% della retina, l'elaborazione neurale associata all'immagine della fovea impiega circa il 50% della corteccia visiva del cervello.

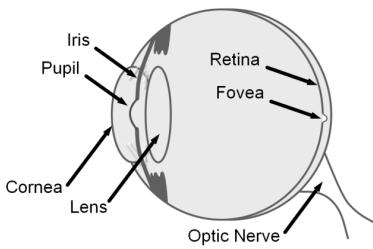


Figura 2.2: Occhio

La vista ha 2 importanti proprietà:

- **Fixation:** gli occhi sono fermi e captano i dettagli visivi dell'ambiente. Possono essere lunghe o brevi, ma in genere durano 200ms.
- **Saccade:** Rapido riposizionamento dell'occhio per fissare un nuovo punto, in genere dura 30-120ms

NOTA: Le ricerche sul tracking dell'occhio sono utilizzate per capire su quale parte dell'interfaccia si focalizza un utente, questo può essere molto utile per capire dove posizionare un adv.

2.1.2 Udito (audition)

L'udito è la rilevazione del suono da parte dell'uomo. Il suono viene trasmesso attraverso l'ambiente sotto forma di onde sonore, fluttuazioni cicliche di pressione in un mezzo come l'aria.

Un modo in cui si creano le onde sonore è quando si sposta un oggetto fisico oppure quest'ultimo viene fatto vibrare creando così fluttuazioni nella pressione dell'aria. Nell'uomo a creare il suono è la laringe (corde vocali), nella gola.

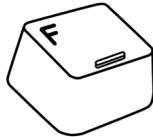
L'**udito** si verifica quando le onde sonore raggiungono l'orecchio umano e stimolano il timpano per creare impulsi nervosi che vengono inviati al cervello. Un singolo suono proveniente da una singola sorgente ha almeno quattro proprietà fisiche: intensità (loudness), frequenza (pitch), timbro e envelope.

- **Loudness:** è analoga all'**intensità**, quantifica il livello di pressione sonora, l'unità di misura è il decibel (dB).
- **Pitch:** è analoga alla frequenza, numero di cicli di un evento che si verificano in un'unità di tempo. Si misura in Hertz (Hz).
- **Timbre:** distingue la struttura armonica dei suoni, è colui che permette di distinguere le varie voci delle persone per esempio.
- **Envelope:** è il modo in cui una nota e le sue armoniche si accumulano nel passare del tempo (da silenziosi a udibili).

2.1.3 Tutto

Questo sistema comprende recettori sensoriali nella pelle, nei muscoli, nelle ossa, nelle articolazioni e negli organi che forniscono informazioni su una serie di fenomeni fisici o ambientali.

Il feedback tattile, in HCI, si riferisce alle informazioni fornite attraverso il sistema sommato sensoriale da una parte del corpo, come ad esempio un dito, quando è a contatto con un oggetto fisico. Tutte le interfacce utente che prevedono il contatto fisico con le mani (o altre parti del corpo) dell'utente includono il feedback tattile.



La linea sotto un tasto della tastiera serve per indicare in che sezione di quest'ultima ci troviamo, è molto utile per la dattilografia.

Press F to pay respect



Il rumore del click del mouse avvisa l'utente che sta eseguendo un'azione.



Quando togliamo la suoneria veniamo avvisati da una vibrazione.

2.1.4 Gusto (taste) e olfatto (smell)

L'olfatto è la capacità di percepire gli odori.

Il gusto (gustation) è la ricezione chimica diretta delle sensazioni di dolce, salato, amaro e acido attraverso le papille gustative della lingua e della cavità orale.

Odori e sapori complessi possono essere costruiti a partire da elementi più semplici, ma i processi percettivi ad oggi rimangono un argomento di ricerca. Sebbene gli esseri umani utilizzino continuamente l'olfatto e il gusto senza sforzo, questi sensi non sono in genere "progettati" nei sistemi. Ci sono alcuni esempi in HCI:

- Brewster et al. (2006): hanno studiato l'olfatto come aiuto nella ricerca di album fotografici digitali. Gli utenti hanno utilizzato due metodi di etichettatura, il testo e l'odore, e poi hanno usato i tag per rispondere a domande sulle foto. Poiché l'olfatto è legato alla memoria, si è ipotizzato che gli indizi olfattivi che i segnali olfattivi potessero aiutare a ricordare le foto. Alla fine, il ricordo con le etichette olfattive è risultato più scarso rispetto a quello con i tag delle parole.
- Bodnar et al. (2004): hanno confrontato modalità olfattive, uditive e visive per notificare agli utenti l'interruzione di un messaggio in arrivo. Anche loro hanno riscontrato prestazioni inferiori con l'olfatto.

L'unione dei processi che trattano gusto e olfatto è detto **flavour**.

2.2 Responders

I movimenti motori di un umano vengono catturati dai responders. Esempio:

- Un dito per scrivere un testo o indicare;
- Piedi per camminare o correre;
- Sopracciglia per alzare le sopracciglia (frown);
- Corde vocali per parlare;

2.2.1 Handedness

Molti utenti sono mancini, destrosi o ambidestri. L' Handedness "exists by degree", significa che qualcosa esiste in varie misure o livelli, anziché essere presente o assente in modo assoluto. Per misurare questo fattore è stato creato un test chiamato "Edinburgh Inventory fo Handedness"

Edinburgh Inventory fo Handedness

| | Left | Right |
|-------------------------|--------------------------|--------------------------|
| 1. Writing | <input type="checkbox"/> | <input type="checkbox"/> |
| 2. Drawing | <input type="checkbox"/> | <input type="checkbox"/> |
| 3. Throwing | <input type="checkbox"/> | <input type="checkbox"/> |
| 4. Scissors | <input type="checkbox"/> | <input type="checkbox"/> |
| 5. Toothbrush | <input type="checkbox"/> | <input type="checkbox"/> |
| 6. Knife (without fork) | <input type="checkbox"/> | <input type="checkbox"/> |
| 7. Spoon | <input type="checkbox"/> | <input type="checkbox"/> |
| 8. Broom (upper hand) | <input type="checkbox"/> | <input type="checkbox"/> |
| 9. Striking a match | <input type="checkbox"/> | <input type="checkbox"/> |
| 10. Opening box (lid) | <input type="checkbox"/> | <input type="checkbox"/> |
| Total (count checks) | <input type="text"/> | <input type="text"/> |
| Difference | <input type="text"/> | <input type="text"/> |
| Cumulative Total | <input type="text"/> | <input type="text"/> |
| RESULT | <input type="text"/> | |

Istruzioni:

Marca le caselle come seguono:

x preferenza

xx preferenza forte

vuoto nessuna preferenza

Punteggio:

Conta il numero di "x" nella colonna left e Right. In comulative total aggiungi il risultato della somma di left e Right. In Difference metti il risultato della sottrazione tra Right e Left. Infine In result inserisci la divisione fra "Difference" e "Comulative Total" e moltiplica per 100.

Interpretazione dei risultati:

da -100 a -40 si è mancini;

da -40 a +40 si è ambidestri;

da +40 a +100 si è destri;

2.2.2 Voce umana

Le corde vocali si comportano da responders. Il suono infatti è creato dalla combinazione del movimento della laringe e dalla pressione polmonare. Esistono due tipi di suoni vocali:

- Speech: da qui nasce la speech recognition
- Non-speech: da qui nasce la signal detection, basata sulla frequenza, l'intensità , la durata, il cambio di direzione di un suono.

2.2.3 NVVI

L'interazione vocale non verbale (NVVI) utilizza vari parametri acustici come l'altezza, il volume o il timbro per rappresentare infine un'informazione (simile al funzionamento del T9). Questa tecnica è particolarmente utile per specificare i parametri analogici.

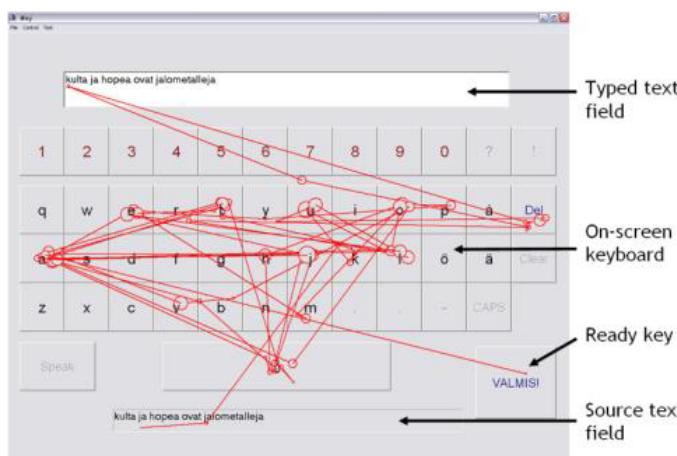
| | Key 1 | Key 2 | Key 3 | Key 4 | BACK |
|-------|-------|-------|-------|-------|------|
| SET 1 | — | — | — | — | — |
| SET 2 | — | — | — | — | — |
| SET 3 | — | — | — | — | — |
| SET 4 | — | — | — | — | — |

Figura 2.3: Alcuni comandi sperimentali del NVVI

2.2.4 Occhio usato come responder

Esiste un controller per catturare i movimenti degli occhi chiamato "double duty":

- Viene utilizzato per sentire e percepire l'ambiente/computer;
- Agisce da controllore tramite la fixation e la saccade.



Il sistema mostrato in figura permette di scrivere tramite l'uso degli occhi. La conferma di una lettera avviene tramite la fixation, e non tramite un blink perché quest'ultimo è involontario e potrebbe quindi confermare una lettera non voluta.

2.2.5 Il cervello

Il cervello è una delle strutture biologiche più complesse che esista. È composta da miliardi di neuroni. Il cervello collega gli input umani(sensori) e gli output umani(responders). Le funzioni del cervello sono le seguenti:

- **Percezione:** è il primo processo che effettua il cervello quando arriva un input esterno. Dalla fine del 19 secolo, la percezione è stata studiata dalla psicofisica. La psicofisica esamina la relazione tra la percezione umana e i fenomeni fisici. Un obiettivo è quello di misurare la differenza appena percettibile (JND) ad uno stimolo. Ad un soggetto umano vengono presentati due stimoli, uno dopo l'altro. Gli stimoli differiscono in una proprietà fisica, come la frequenza o l'intensità, e al soggetto viene chiesto se gli stimoli sono uguali o diversi. Questo può partire a vedere delle ambiguità in un'immagine oppure a subire un'illusione ottica.
- **Cognizione:** è il processo che permette di pensare, ragionare e prendere delle decisioni. Viene studiata in molti campi: neurologia, linguistica, antropologia.

I **fenomeni sensoriali(sensory phenomena)** come il suono e la luce sono facili da studiare perché esistono nel mondo fisico, esistono infatti numerosi strumenti per misurarli.

La **cognizione (Cognitive phenomena)** si verifica all'interno del cervello umano, quindi studiarla è molto difficile. Ad esempio, non è possibile misurare direttamente il tempo impiegato da un essere umano per prendere una decisione.

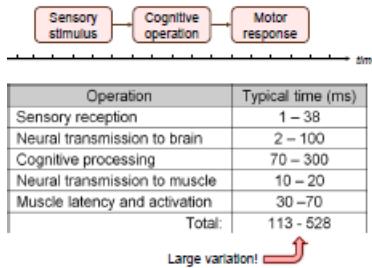


Figura 2.4: Tempi di risposta

- **Memoria:** in grado di contenere immagazzinare una grande quantità di dati. Esistono due tipi di memoria:
 - **Lungo termine:**
 - * **Declarative/explicit area:** informazioni su eventi nel tempo e oggetti nel mondo esterno.
 - * **Implicit/procedural area:** informazioni su come usare gli oggetti e come fare le cose.
 - **Breve termine:** chiamata anche working memory (permette di mantenere attivamente e manipolare informazioni durante il processo di pensiero). Le informazioni sono sempre attive e disponibili ad un accesso. La quantità di memoria di lavoro è piccola, circa 7 (± 2) unità o chunks.
Di seguito sono mostrate le sequenze di caratteri che una persona può ricordare: 7 circa il 50%. 5 circa il 90% e 9 circa il 20%.
- **Linguaggio:** è la capacità che permette agli esseri umani di comunicare. La possibilità di comunicare è disponibile senza sforzi, mentre la capacità di scrivere richiede un considerevole sforzo. Nell'**HCI** interesse primario è la scrittura, per la creazione di testi.

L'umanità è definita dal linguaggio; ma la civiltà è definita dalla scrittura.

Un modo per caratterizzare il testo scritto è un **corpus**. È una grande collezione di testo. Un corpus si può ridurre in una lista di frequenze delle parole.

| Word Rank | English | French | German | Finnish | SMS English | SMS Pinyin |
|-----------|---------|-----------|------------|----------|-------------|------------|
| 1 | the | de | der | ja | u | wo (我) |
| 2 | of | la | die | on | i | ni (你) |
| 3 | and | et | und | ei | to | le (了) |
| 4 | a | le | in | että | me | de (的) |
| 5 | in | à | den | oli | at | bu (不) |
| ... | ... | ... | ... | ... | ... | ... |
| 1000 | top | ceci | konkurrenz | muista | ps | jiu (舅) |
| 1001 | truth | mari | stieg | paikalla | quit | tie (贴) |
| 1002 | balance | solution | notwendig | varaa | rice | ji (即) |
| 1003 | heard | expliquer | sogenannte | vie | sailing | jiao (角) |
| 1004 | speech | pluie | fahren | seuran | sale | ku (裤) |
| ... | ... | ... | ... | ... | ... | ... |

Figura 2.5: Frequenza delle parole

Queste frequenze possono essere utilizzate per anticipare le lettere, delle parole o intere frasi.

SMS Shorthand è un modo di scrivere le parole in modo abbreviato, questo può portare ad un'incomprensione del testo.

2.3 Human Performance

Gli esseri umani utilizzano sensori, rispondono e il cervello per fare cosa. Quando queste tre componenti lavorano insieme per raggiungere un obiettivo, nasce la performance umana (es: ricercare sul web, piegare (folding) i panni, ecc).

2.3.1 Speed-accuracy Trade-off

La human performance ha dei trade-off all'aumentare della velocità:

- Più si va veloce e più gli errori aumentano;
- Se si rallenta e la precisione migliora;
- La ricerca **HCI** su una nuova interfaccia o tecnica di interazione deve considerare sia la velocità nell'esecuzione dei compiti (raggiungimento dell'obiettivo!) sia la relativa accuratezza.

2.3.2 Human Diversity

La human performance è influenzata da molti fattori: ogni persona però è differente (età, genere, skill, ecc), le condizioni dell'ambiente in cui vengono effettuati i test di performance, la presenza di compiti secondari.

La human diversity e la human performance possono essere mostrati su una distribuzione gaussiana (normale).

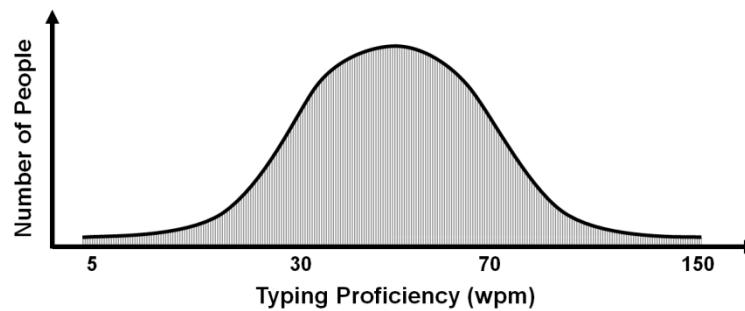


Figura 2.6: Distribuzione human diversity e human performance

2.3.3 Reaction Time

Una delle manifestazioni più primitive della performance umana è il semplice tempo di reazione.

Definizione: il ritardo tra il verificarsi di un singolo stimolo e l'inizio di una risposta ad esso assegnata. **Esempio:** pressione di un pulsante in risposta all'insorgenza di uno stimolo luminoso.

Il tempo di risposta può subire un delay in base a vari stimoli sensoriali:

- Udito: 150 ms
- Visivo: 200 ms
- Odore: 300 ms
- Dolore: 700 ms

Alcuni esperimenti per il tempo di risposta:

- **Simple Reaction Time:** viene mostrato un quadrato, appena si illumina si deve premere un qualsiasi tasto;
- **Physical Matching:** vengono mostrate due parole se sono uguali si deve premere J altrimenti k;
- **Name Matching:** come il Physical Matching solo che stavolta le scritte hanno un formato di font sempre differente;
- **Class Matching:** come i due precedenti solo che si verificano che le lettere siano in coppia con le lettere e i numeri con i numeri;
- **Visual Search:** variante del physical matching, si controlla se una lettera è presente in un insieme di caratteri, l'insieme può essere di 1,2,4,8,16 o 32 elementi.

2.3.4 Skilled Behaviour

Per molte attività, le prestazioni umane migliorano considerevolmente e continuamente con la pratica (Nota: Miglioramento molto limitato con la pratica nelle semplici attività di tempo di reazione). In questi compiti, c'è interesse a studiare la progressione dell'apprendimento e le prestazioni ottenute in base alla quantità di pratica.

- Sensory-motor skill (es: gaming);
- Mental skill (es: scacchi, programmare)

Interaction Elements

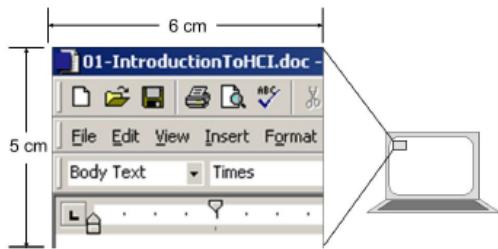
In questo capitolo si parla della parte "computer" del HFM.

3.1 Hard e Soft controls

Prima che i P.C. entrassero nella vita di tutti, i display e i controlli tendevano a essere dispositivi fisici (joystick, pulsanti, tasti, ecc), questi erano chiamati **Hard controls**. La nascita dei P.C. e lo sviluppo dei software e successivamente la nascita dei display grafici ha portato alla nascita dei **Soft controls** cioè delle interfacce grafiche inseriti direttamente nei software. I soft control sono facilmente malleabili. Esiste una differenza fra display e soft control:

- Un soft control è reindirizzato su un display mediante l'uso di un software, ha una dimensione, una forma o un'apparenza che cambia in base alle informazioni che decide l'utente;
- Un display invece è un elemento mostrato all'utente.

NOTA: Un soft control può possedere delle proprietà di un display, un esempio sono le toolbar o widget.



Esempio di malleabilità di un soft control, in uno spazio di 30 cm^2 abbiamo circa una 20-tina di soft control, esiste anche la possibilità che cliccando su di un bottone la GUI cambi completamente.

3.2 Relazioni tra control e display

La relazione che esiste tra control e display prende il nome di **mapping**, e descrive il modo di come il funzionamento di un controller è mappato su di un display (es: lo spostamento del mouse da parte di un utente fa sì che il cursore sul display si muova).

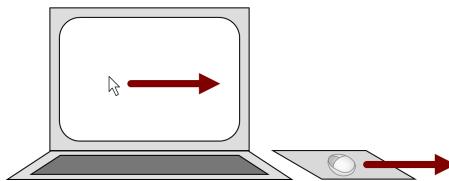
Esistono 3 tipi di mapping:

- **Spaziale:** descrive come un'azione del controller nel mondo reale viene mappata nel mondo digitale per esempio su di un display;
- **Dinamico:** descrive come il controller influenza la velocità di risposta;
- **Fisico:** descrive se la risposta è un movimento/forza del controllore

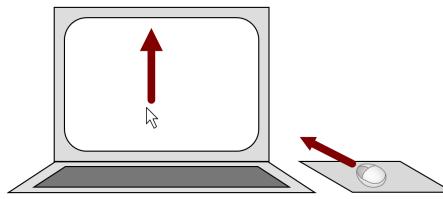
3.3 Mapping Spaziale

Il mapping spaziale può essere di due tipologie:

- **Spatial congruence:** è una corrispondenza 1 a 1 tra l'input del controller e l'output sul display; è un movimento naturale;



- **Spatial transformation:** il controller può effettuare un movimento in due dimensioni, normalmente quest'ultimo deve essere mappato su un display a due dimensioni; questo non è un movimento naturale ma è un movimento che viene appreso.



3.3.1 Axis Labeling

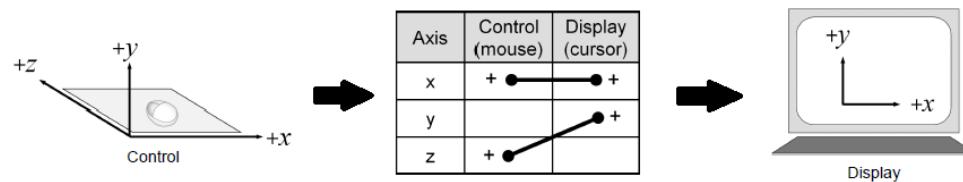


Figura 3.1: Un movimento sull'asse x è mappato sull'asse x, uno nell'asse z in y

3.3.2 Third Tier

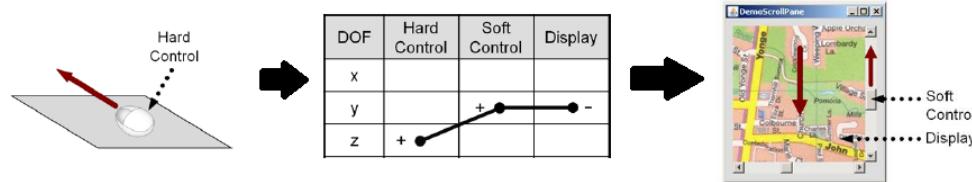


Figura 3.2: Un movimento sull'asse < è mappato sull'asse y con uno spostamento negativo.

3.3.3 Trasformazione spaziale: 3D

È una trasformazione spaziale, in un ambiente 3D abbiamo i così detti 6 gradi di libertà (DOF). Nella **DOF** ogni parametro può essere manipolato in maniera indipendente dagli altri. I parametri presenti che possono essere manipolati sono 6:

- 3 parametri per la posizione nello spazio: x y z;
- 3 parametri per l'orientamento nello spazio: $\theta_x \theta_y \theta_z$

Le rotazioni possibili sui 3 assi in aeronautica hanno dei nomi:

- θ_x : **Pitch** (beccheggio o impennata)
- θ_y : **Yaw** (imbardata o serpeggiamento)
- θ_z : **Roll** (rollio o coricamento)

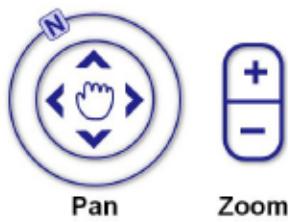


Figura 3.3: Grado di libertà un pò troppo oltre

Come capire se la rotazione è positiva o negativa? Si può utilizzare la regola del pollice o meglio la regola del pugno chiuso. Se la mano ruota in avanti è positiva se ruota all'indietro è negativa.

3D in Google street view - Panning in Google StreetView

Un sottosistema del 6 DOF è utilizzato in google street view, e rappresenta un sistema di tipo learned



Mapping dei comandi in Google StreetView

3.3.4 Trasformazione spaziale: CD Gain (CD-Ratio)

È anch'essa una trasformazione spaziale, e quantifica la quantità di movimento mostrata nel display a seconda di quanto si è mosso il controller. Esempio 2 movimenti di 2 centimetri sul controller portano ad un singolo movimento di 4 cm sul display.

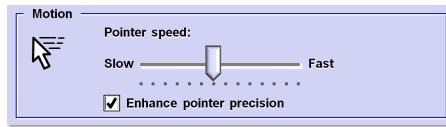


Figura 3.4: Esempio di CD gain

Questo sistema ha dei problemi tutti voltati ad ottimizzare le performance degli utenti, i problemi che riscontriamo sono:

- **Trade off per l'accuratezza della velocità:** se si riduce il tempo di posizionamento si incrementano gli errori;
- **Relazione opposta tra gross(denso) e fine (sottile)**

Latenza : La latenza è il ritardo tra un'azione in input e la corrispondente risposta sul display. Di solito è trascurabile nei sistemi interattivi, mentre ad esempio in altri casi è tangibile come ad esempio nel caso delle manipolazioni remote o nella VR.

3.4 Propeery sensed e control

controller rilevano le interazioni e le racchiudono in property sensed cioè dati che vengono trasmessi al computer per essere processati. Le property sensed tradotte sono le tecniche di interazione dell'utente e possono essere:

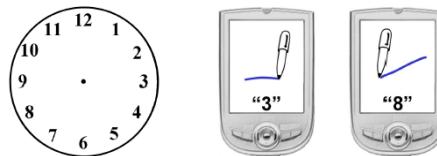
- **Position:** come ad esempio tablet, touchpad, touchscreen;
- **Displacement (spostamento):** come ad esempio mouse, joystick;
- **Force:** come ad esempio pointing stick e joystick.

Il cambiamento prodotto detto order control nella vista del display può essere di due tipi:

- **Position** del cursore;
- **Velocità** del cursore.

3.5 Metafora dell'orologio

È stata utilizzata per poter scrivere i numeri sui PDA, e si basa sul tracciare una linea verso la **DIREZIONE** di una lancetta dell'orologio.



Questo sistema è utilizzato dalle persone cieche per potersi spostare, o anche dall'aeronautica. Per le persone cieche funziona nel seguente modo. Un dispositivo indica in formato audio in quale punto della stanza si trova un ostacolo.

3.6 Modalità (Mode)

Le modalità sono una disposizione e ricollocamento di più funzionalità su di una combinazione di tasti.

La tastiera del computer ha molte modalità: 100 tasti + shift, ctrl, alt generano circa 800 variazioni di tasti.

Un primo esempio è il pulsante navi key(introdotto per la prima volta nel Nokia 3210), incorporava ben 15 funzionalità. Un altro esempio con un'idea simile sono i monito LCD contemporanei, ogni pulsante incorpora un sotto-menù software con diverse modalità.

3.6.1 Mode switching

Un esempio di cambio di modalità cambiando la grafica del programma è PowerPoint, la presenza in una determinata modalità è data da un'ombreggiatura del background. Le modalità possibili sono 5.

Un altro esempio di mode switching sono le modalità di un orologio sportivo, in cui un singolo pulsante attiva diverse modalità (per via di questo funzionamento spesso è richiesto l'uso di un manuale).

3.6.2 Mode visibility

La mode visibility è utilizzata per dare un feedback grafico all'utente nell'entrata in una nuova modalità. L'editor *vi* in unix è un classico esempio di un sistema senza mode visibility. Un altro esempio è il tasto insert di word, che se premuto attiva la modalità overtype, il problema è che non viene segnalato in nessun modo la sua attivazione.

Gradi di libertà: esistono diversi gradi di libertà (2, 4), ecc.. oppure possono essere separati come per esempio nel sistema: Etch-A-Sketch in cui abbiamo 1 DOF e due controller per spostarsi nell'asse x e y.

3.7 Mobile context

Con l'evoluzione della tecnologia e la nascita di nuovi dispositivi è cambiato anche il concetto di input che si divide in:

1. **Indirect input:** come ad esempio mouse e touchpad in cui la risposta è fornita (spazialmente) lontana dal dispositivo utilizzato per fornire l'input;
2. **Direct input:** come ad esempio dita, display o touch in cui la risposta è fornita (spazialmente) nello stesso punto usato per l'input.

3.7.1 Touch screen

Uno degli esempi principali di direct input è il touch screen. Sono diventati molto fondamentali per via della loro grande direct manipulation.

Per essi, infatti, non vi è bisogno di utilizzare un cursore o di effettuare un mapping control/display. Uno dei problemi però principale fa riferimento al sistema utilizzato per il touch: la selezione precisa dei pixel è difficile. Tale sistema di solito è composto dalle dita delle persone.

Si parla di **FAT FINGER PROBLEM** quando il dito è più grande del tasto da premere. A tale problema è stata fornita una soluzione da Potter nel 1988 che prevedeva l'utilizzo dell'offset cursor o anche detta take-off selection. Tale tecnica prevedeva l'uso di un mirino che appare sopra la punta delle dita che toccano il display.

Una differenza tra mouse e touch (e quindi tra direct e indirect) è che il touch è sensibile anche al multitouch. Bisogna inoltre sottolineare che per multitouch non si intende solo l'utilizzo di due dita: basti pensare all'applicazione Piano keyboard che riconosce la pressione di 10 tasti contemporaneamente.

3.8 Accelerometro

I dispositivi moderni oltre all'utilizzo del multitouch possiedono anche altri componenti aggiuntivi come ad esempio l'accelerometro. Tale componente permette di rilevare l'angolo e l'inclinazione (TILT) del dispositivo. Proprio il tilt è divenuto una delle principale tecniche di interazione nel mobile computing.

L'accelerometro è anche utilizzato nella visualizzazione spaziale (aware display) cioè il movimento della vista sul display muovendo il dispositivo : l'accelerometro rileva il movimento del dispositivo e genera un movimento della vista sul display.

Fondazioni scientifiche

La definizione di ricerca (research) è la seguente:
"insieme di indagini o investigazioni che mirano alla scoperta e all'interpretazione dei fatti, alla revisione di teorie o leggi accettate alla luce dei nuovi fatti."

Ad esempio progettare o condurre uno studio per verificare se una nuova tecnica di interazione è migliore di una tecnica di interazione esistente.

4.1 Sperimentazione

La sperimentazione è una delle attività centrali della ricerca HCI. Un modo alternativo per chiamare gli esperimenti è usare i termini user study.

È preferibile usare la **metodologia formale e standardizzata** poiché porta:

- Maggiore consistenza;
- Facili review e confronti tra diversi "user study".

4.2 Pubblicazione

Il fine ultimo di una ricerca è la pubblicazione, è uno dei passi essenziali.

Si parla di **peer review** (revisione paritaria) quando una ricerca è revisionata da ricercatori che svolgono una ricerca simile.

Queste ricerche possono essere pubblicate per far conoscere a tutti i risultati.

In altri casi si può presentare un **brevetto (patent)**, che descrive:

- altri lavori correlati;
- come la ricerca effettuata è in grado di risolvere un problema;
- qual'è il miglior modo per implementarlo.

Le citazioni connettono tra di loro le varie ricerche. Il numero di citazioni di una ricerca è un'indicazione dell'influenza di essa. Esse sono usate per valutare le sedi di pubblicazione infatti si parla di impact factor.

Grazie all'invenzione della DL (Digital Library) come ad esempio Google Scholar o Scopus il calcolo delle citazioni è molto più semplice : è automatico. Inoltre tali librerie dividono direttamente le ricerche per giornali, riviste, ricercatori, ecc.

4.3 H-index

L'indice h (o h-index) è un indicatore bibliometrico che misura la produttività e l'impatto di un ricercatore accademico, basandosi sul numero di citazioni che i suoi articoli scientifici hanno ricevuto. L'indice h viene calcolato contando il numero di articoli (n) pubblicati da un autore che hanno ricevuto almeno n citazioni ciascuno.

Ad esempio, se un autore ha pubblicato 10 articoli, di cui 5 hanno ricevuto almeno 5 citazioni ciascuno, l'indice h dell'autore sarebbe 5.

Il modo migliore per semplificare questo calcolo è quello di ordinare in ordine decrescente di citazioni ricevuti gli articoli di un ricercatore, e per ogni articolo vedere a quale riga si trova e controllare se è minore delle citazioni relative all'articolo per quella riga. Quando si arriva in un caso in cui il numero di riga è maggiore del numero di citazioni allora h-index è il numero di righe totali fino a quel punto.

4.4 Digital library

Le DL per computer science più popolari sono:

- Google scholar per tutti i campi della scienza;
- ACM DL per computer science;
- IEEE Explorer per l'ingegneria;
- Science direct per differenti campi della scienza.

4.5 Valutazioni giornali

I giornali possono essere valutati per i seguenti indici:

- **IF (Impact Factor)**: numero medio annuale di citazioni di articoli recenti pubblicati da quel giornale;
- **SJR (Scimago Journal Risks)**: molto simile al precedente ma le citazioni sono ponderate in base al prestigio;
- **SNIP (Source Normalized Impact per Paper)** citazioni ricevute VS citazioni previste in base al campo del paper

Le **HCI Venues** (sedi) possono essere:

- **Conferenze** come per CHI, UIST, IUI etc
- **Giornali** come ad esempio HCI, TOCHI, Elsevier, Oxford Journal

Per controllare la valutazione di una conferenza andare su: GII-GRIN-SCIE (GGS) Conference Rating

Per controllare la valutazione di un giornale andare su: Scopus

4.6 Riproducibilità di una ricerca

Un'altra caratteristica delle ricerche è che esse possono essere riproducibili (se non lo sono allora sono inutili). In particolare un livello elevato di riproducibilità è essenziale. Ma come facciamo ad avere un alto livello di riproducibilità? Seguendo una metodologia standardizzata. Per calcolare il livello di riproducibilità si utilizzano:

- **Precision** che indica la frazione di documenti rilevanti:

$$\text{precision} = \frac{|\{\text{documenti rilevanti}\} \cap \{\text{documenti recuperati (retrieved)}\}|}{|\{\text{documenti recuperati}\}|} \quad (4.1)$$

- **Recall** indica la frazione di documenti rilevanti recuperati con successo

$$\text{recall} = \frac{|\{\text{documenti rilevanti}\} \cap \{\text{documenti recuperati (retrieved)}\}|}{|\{\text{documenti rilevanti}\}|} \quad (4.2)$$

Molto spesso i ricercatori sono a contatto con ingegneri e progettisti ed ognuno di loro ha competenze differenti. I progettisti e gli ingegneri hanno perlopiù competenze progettuali e ingegneristiche.

| Study (1 st author) | Number of Keys ^a | Direct/ Indirect | Scanning | Number of Participants | Speed ^b (wpm) | Notes |
|-----------------------------------|--------------------------------|---------------------|----------|---------------------------|-----------------------------|---|
| Bellman [2] | 5 | Indirect | No | 11 | 11 | 4 cursors keys + SELECT key. Error rates not reported. No error correction method. |
| Dunlop [4] | 4 | Direct | No | 12 | 8.90 | 4 letter keys + SPACE key. Error rates reported as "very low." |
| Dunlop [5] | 4 | Direct | No | 20 | 12 | 4 letter keys + 1 key for SPACE/NEXT. Error rates not reported. No error correction method. |
| Tanaka-Ishii [25] | 3 | Direct | No | 8 | 12+ | 4 letters keys + 4 keys for editing, and selecting. 5 hours training. Error rates not reported. Errors corrected using CLEAR key. |
| Gong [7] | 3 | Direct | No | 32 | 8.01 | 3 letter keys + two additional keys. Error rate = 2.1%. Errors corrected using DELETE key. |
| MacKenzie [16] | 3 | Indirect | No | 10 | 9.61 | 2 cursor keys + SELECT key. Error rate = 2.2%. No error correction method. |
| Baljko [1] | 2 | Indirect | Yes | 12 | 3.08 | 1 SELECT key + BACKSPACE key. 43 virtual keys. RC scanning. Same phrase entered 4 times. Error rate = 18.5%. Scanning interval = 750 ms. |
| Simpson [24] | 1 | Indirect | Yes | 4 | 4.48 | 1 SELECT key. 26 virtual keys. RC scanning. Excluded trials with selection errors or missed selections. No error correction. Scanning interval = 525 ms at end of study. |
| Koester [10] | 1 | Indirect | Yes | 3 | 7.2 | 1 SELECT key. 33 virtual keys. RC scanning with word prediction. Dictionary size not given. Virtual BACKSPACE key. 10 blocks of trials. Error rates not reported. Included trials with selection errors or missed selections. Fastest participant: 8.4 wpm. |

^a For "direct" entry, the value is the number of letter keys. For "indirect" entry, the value is the total number of keys.
^b The entry speed cited is the highest of the values reported in each source, taken from the last block if multiple blocks.

Figura 4.1: Come organizzare i documenti recuperati

4.7 Ricerca vs Ingegneria vs Design

I ricercatori lavorano spesso a stretto contatto con ingegneri e progettisti, ma i ruoli sono differenti. Molte volte il predominare di un ruolo o un altro può portare a problemi. Per esempio il voler creare interfacce gradevoli da guardare può portare alla creazione per esempio di un touchpad come segue:



Molto bello da vedere, ma poco funzionale perché non si vedono bene i confini del touchpad

Bisogna ricordare quindi che l'ingegneria e il design riguardano i prodotti. La ricerca invece è strettamente focalizzata sul risolvere un problema, tramite l'uso di domande e procedimenti incrementali.

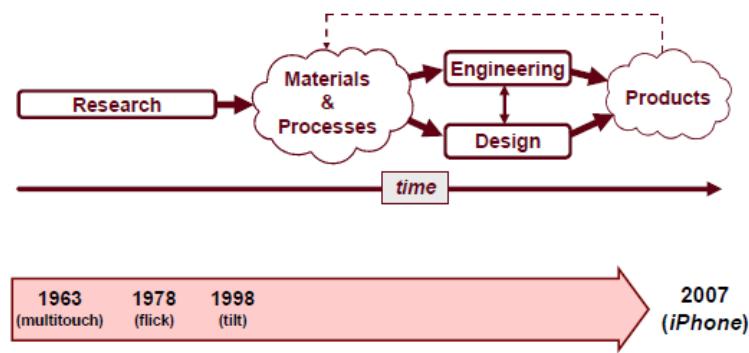


Figura 4.2: Timeline ricerca-ingegneria-design

4.8 Empirical Research

Una ricerca si dice empirica se è originata o basata sull'osservazione o sull'esperienza senza tener conto del sistema o delle teorie cioè senza preconcetti (ad esempio le ricerche di Copernico).

Una ricerca può essere qualitative o quantitative:

- **qualitative:** si riferisce a informazioni non numeriche o dati basati su opinioni o percezioni soggettive. Ad esempio, si potrebbe esaminare la facilità d'uso di un sito web chiedendo aiutando gli utenti a completare delle attività e chiedendo loro feedback sul processo. Queste informazioni sono qualitative in quanto sono basate sulle opinioni degli utenti e non possono essere facilmente quantificate.
- **quantitative:** si riferisce a informazioni numeriche o dati misurabili e oggettivi. Ad esempio, si potrebbe esaminare il tempo necessario per completare un'azione su un sito web e registrare i dati in modo da poterli analizzare statisticamente. Questi dati sono quantitativi in quanto possono essere misurati e quantificati.

I metodi empirici di una ricerca possono essere di 3 tipologie differenti:

1. **Observational**
2. **Experimental**
3. **Correlational**

4.8.1 Observational empirical method

Alcuni esempi possono essere interviste, casi studio, focus group, investigazioni, indagini sul campo, racconti. È essenziale osservare come l'essere umano interagisce con i computer o più in generale con la tecnologia. L'approccio è qualitativo piuttosto che quantitativo. Il comportamento è analizzato direttamente nel suo ambiente naturale e non si crea nessun ambiente ad hoc. Questo metodo ha l'obiettivo di determinare e spiegare le ragioni che stanno alla base del comportamento umano.

Questo metodo ha **Alta rilevanza (High relevance)** comportamenti studiati in un ambiente naturale, ma **Bassa precisione (Low precision)** manca il controllo disponibile in un laboratorio.

4.8.2 Experimental Method

Anche conosciuto come metodo scientifico Si tratti di esperimenti controllati e condotti in laboratori specializzati . La **rilevanza** è **bassa** poiché l'ambiente è artificiale ma la **precisione** è **alta** poiché i comportamenti estranei possono esser tenuti facilmente sotto controllo. Ci sono due variabili fondamentali:

- **Manipulated variable / factor** (variabile indipendente): Si espongono i partecipanti a diverse configurazioni dell'interfaccia o della tecnica di interazione;
- **Response variable** (variabile dipendente): è una proprietà del comportamento umano che è osservabile, quantificabile e misurabile. Sono tutte quelle cose che dipendono dalle caratteristiche dell'interfaccia (ad esempio il tempo impiegato, il tipo di task ecc);

Bisogna tener presente che: “cambiamenti delle variabili indipendenti comportano un cambiamento nelle variabili dipendenti”.

4.8.3 Correlational Method

È quel metodo in cui si osservano le relazioni tra le variabili. Si tratta di:

- **Fare osservazioni**
- **Raccogliere dati**

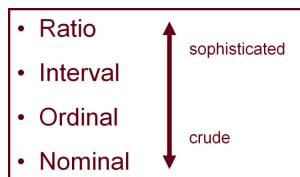
È un metodo non sperimentale che consiste nel fare interviste, sondaggi, questionari. Rilevanza e precisione sono bilanciati.

4.9 Osservare e misurare

Alla base della ricerca empirica vi sono i concetti di osservazione e misurazione. Si ha che:

- L'osservazione è il punto di inizio ed è fatta da umani o apparati . L'osservazione può anche essere manuale fatta cioè tramite notebook, screen, fotografie, video etc.
- La misurazione mentre permette di ottenere l'evidenza empirica

La scala di misura va dalla misura più sofisticata a quella più grezza (crude) ed è la seguente:



4.9.1 Nominal Data

Sono i dati categorizzati come ad esempio il numero della piazza, il codice fiscale, la classificazione del lavoro. In alcuni casi è possibile che venga associato un codice o un attributo; ad esempio 1 per maschio e 2 per femmina . Si può facilmente comprendere che non possiamo effettuare analisi statistiche.

4.9.2 Ordinal Data

Sono i dati che possiedono un ordine o un rank associato ad un attributo. Sono possibili i confronti fra essi. Gli attributi che vengono misurati non sono altro che caratteristiche o interessi. Ad esempio supponiamo di considerare l'esperimento che consiste nell'invio di una mail. Si parla di dati ordinali se abbiamo le seguenti risposte:

- 1) nessuna Email;
- 2) invio di 1-5 email;
- 3) invio di 6-10 email;
- 4) etc;

4.9.3 Radio Data

Sono quei dati più sofisticati e che permettono tutti i calcoli (somma, comparare, divisione, ecc). Hanno uno zero assoluto. NellHCI una delle scale più utilizzate è il tempo (es: 20 parole digitate al minuto (wpm)).

4.9.4 Interval Data

Sono dati in cui si hanno degli intervalli come ad esempio i dati relativi alle temperature . Per tale tipologia di dati è possibile effettuare dei calcoli statistici come ad esempio determinare qual è la temperatura media. Un esempio di tali dati sono quelli misurati mediante il metodo della scala di Likert. Non è possibile fare rapporti (es: 10° C non è il doppio 5°C, è dovuto alla mancanza dello 0 assoluto).

| Please indicate your level of agreement with the following statements. | | | | | |
|--|----------------------|--------------------|---------|-----------------|-------------------|
| | Strongly disagree | Mildly disagree | Neutral | Mildly agree | Strongly agree |
| It is safe to talk on a mobile phone while driving. | 1 | 2 | 3 | 4 | 5 |
| It is safe to read a text message on a mobile phone while driving. | 1 | 2 | 3 | 4 | 5 |
| It is safe to compose a text message on a mobile phone while driving. | 1 | 2 | 3 | 4 | 5 |

4.9.5 Quale funzioni calcolabili si possono usare sulle variabili?

| Ok to compute | Nominal | Ordinal | Interval | Ratio |
|--|---------|---------|----------|-------|
| Frequency distribution | Yes | Yes | Yes | Yes |
| Median and percentiles | No | Yes | Yes | Yes |
| add or subtract | No | No | Yes | Yes |
| mean, standard deviation, standard error of the mean | No | No | Yes | Yes |
| ratio, or coefficient of variation | No | No | No | Yes |

Tabella 4.1: Funzioni calcolabili per il tipo di variabile

4.9.6 Internal or External Validity

Altre proprietà importanti degli esperimenti sono:

- **Internal validity:** con cui si intende la misura per cui gli effetti osservati dipendono dalle condizioni dei test di prova. Ad esempio se confrontiamo la velocità delle due tecniche di text entry vogliamo esser sicuri che la differenza osservata faccia riferimento realmente alla differenza di velocità e che non sia associata ad altri fattori come ad esempio incapacità di un partecipante ad utilizzare un metodo piuttosto che un altro (può essere influenzata infatti da fattori esterni, es: un ambiente rumoroso (noisy) può deconcentrare i partecipanti).
- **External validity:** con cui si intende la misura per cui i risultati sono generalizzabili alle persone o alle situazioni. Le persone rappresentano la popolazione che ha svolto gli esperimenti mentre per situazioni si intende il contesto in cui sarà usata la tecnica provata.

Trade off fra le due proprietà

C'è tensione tra validità interna ed esterna. Più l'ambiente di prova e le procedure sperimentali sono "rilassate" (per imitare le situazioni del mondo reale), più l'esperimento è suscettibile a fonti incontrollate di variazione, come ponderazione, distrazioni, giocherellare o attività secondarie.

Valutazione ricerca

È preferibile fare una valutazione comparativa fra più casi di studio piuttosto che valutare il proprio singolo caso. Si richiede il confronto, e deve esserci almeno una variabile indipendente con almeno due livelli misurabili.

Come fare un questionario?

Una parte importante della ricerca sono le domande da porre. Queste devono interrogare su qualcosa di misurabile:

- Very weak: Is the new technique any good?
- Weak: Is the new technique better than QSK?
- Better: Is the new technique faster than QSK?
- Better still: Is the measured entry speed (in words per minute) higher for the new technique than for QSK after one hour of use?

Design HCI experiment

Progettare un esperimento significa scegliere i partecipanti, quale task far eseguire, in quale ordine, quali sono le azioni da eseguire, le variabili (dipendenti e indipendenti) e infine i dati da testare.

Abbiamo 3 tipi di design :

- **Within-subjects:** tutti i partecipanti vengono sottoposti a tutte le condizioni dell'esperimento;
- **Between-Subjects:** i partecipanti vengono divisi in gruppi, e ogni gruppo riceve un diverso tipo di dato su cui effettuare l'esperimento. I partecipanti testano solo una condizione
- **Counterbalancing:** è una tecnica utilizzata in alcune ricerche scientifiche per evitare che l'ordine in cui vengono somministrate le condizioni di un esperimento influenzi i risultati. In pratica, si cerca di bilanciare gli effetti dell'ordine di somministrazione delle condizioni su tutti i partecipanti allo studio. In questo modo, si può avere maggiore sicurezza sui risultati ottenuti e si evitano eventuali distorsioni dovute all'ordine di somministrazione delle condizioni.

L'experiment design è una parte della metodologia e rappresenta il modo in cui si progetta un certo test. Possiamo dire che: "La scienza è il metodo. Tutto il resto è un commento"

APA

American Psychological Association (APA) è l'organizzazione predominante che promuove la ricerca nella psicologia, studia come migliorare le condizioni di ricerca e l'applicazione dei risultati della ricerca. La prima pubblicazione dell'APA è del 1929, descriveva il processo di scrittura e implicitamente, la metodologia per gli esperimenti con partecipanti umani. Questo articolo è raccomandato dalla maggior parte dei giornali nellHCI.

5.1 Ethics Approval

Uno dei passi principali di HCI è l'approvazione etica (ethics approval).

Dato che HCI ha a che fare con gli esseri umani:

I ricercatori devono rispettare la sicurezza, il benessere e la dignità dei partecipanti umani alla loro ricerca e trattarli in modo equo.

Tale processo è controllato da un ente che si occupa di supervisionare la ricerca (HRPC).

Iniziare con la progettazione (Design) degli esperimenti

1. Formulare una domanda per la ricerca:

Un'attività può essere eseguita più rapidamente con la mia nuova interfaccia rispetto a un'interfaccia esistente?

2. Si ci chiede quali sono le variabili dell'esperimento?

Il tempo di completamento del task è la variabile dipendente, mentre la nuova e la vecchia interfaccia sono quelle indipendenti

5.1.1 Variabile indipendente

Una variabile indipendente è una caratteristica o una circostanza che può esser manipolata per provocare un cambiamento nella risposta umana. Genericamente è un fattore che non dipende dal comportamento dei partecipanti, ma se cambia può provocare un cambiamento.

Esempio: Consideriamo una stanza. Il colore della stanza è una variabile indipendente poiché non dipende dal comportamento delle persone. Però cambiare il colore di una stanza potrebbe far cambiare le impressioni alle persone presenti nella stanza. Quando parliamo di circostanze intendiamo la presenza di rumore di fondo, l'illuminazione, e livello di vibrazione (frequenza delle oscillazioni di un oggetto o di un sistema), ecc...

Quando si parla di caratteristiche si fa spesso riferimento alle caratteristiche umane come ad esempio: l'età, il sesso, S.O. preferito, prima lingua, ecc...

Queste però sono variabili che non possono esser manipolate allo stesso modo di una circostanza.

Infatti il termine **indipendente** fa riferimento al fatto che la variabile non dipende dal comportamento dei partecipanti. Queste variabili prendono anche il nome di fattori.

Test conditions

Le "Test conditions" sono le specifiche condizioni sotto cui viene condotto uno studio o un esperimento per valutare l'effetto delle variabili indipendenti sulla variabile dipendente. Una variabile indipendente (IV) ha quindi due livelli che sono appunto le test conditions. Quindi abbiamo i fattori (IV) e i livelli (Test conditions).

| Factor (IV) | Levels (test conditions) |
|------------------|----------------------------|
| Device | mouse, trackball, joystick |
| Feedback mode | audio, tactile, none |
| Task | pointing, dragging |
| Visualization | 2d, 3d, animated |
| Search interface | Google, custom |

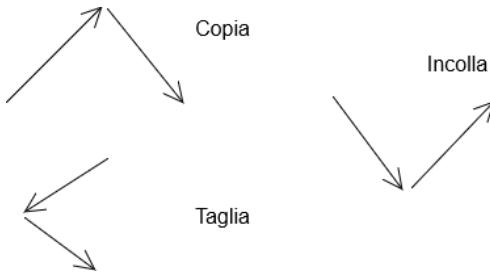
Ogni esperimento deve avere almeno una variabile indipendente. Generalmente abbiamo m variabili e n livelli, da qui otteniamo il **factorial design** uguale a $m \times n$.

Esempio: TouchTap

È un implementazione delle gesture sulla tastiera per migliorare le azioni di editing del testo. Migliora le vecchie tecniche utilizzando le gesture per effettuare operazioni di editing e permette l'utilizzo dello smartphone con la "bi-manual posture". Per entrare in questa modalità basta tenere premuto un dito sulla tastiera, e fare le gesture con un altro dito.

Le funzionalità offerte sono:

- Muovere il cursore muovendo il dito a destra o sinistra;
- Selezionare del testo;
- Auto-repeat così da selezionare più testo;
- Clipboard tramite delle gesture:



Come design si è utilizzato il **within-subjects**.

Le IV erano:

- Input method : TouchTap e il widget base per scrivere;
- Font size : 1.75 , 3.25 , 4.75.

Il factorial design è quindi 2x3, ci sono state quindi 6 test da testare.

Quante variabili indipendenti usare?

Sappiamo che ogni esperimento debba avere almeno una IV, ma è possibile averne 2, 3 o più di una. Ma il numero di "effetti" da controllare poi aumenta incredibilmente:

| Independent Variables | Effects | | | | | Total |
|-----------------------|---------|-------|-------|-------|-------|-------|
| | Main | 2-way | 3-way | 4-way | 5-way | |
| 1 | 1 | - | - | - | - | 1 |
| 2 | 2 | 1 | - | - | - | 3 |
| 3 | 3 | 3 | 1 | - | - | 7 |
| 4 | 4 | 6 | 3 | 1 | - | 14 |
| 5 | 5 | 10 | 6 | 3 | 1 | 25 |

5.1.2 Variabile dipendente (DP)

Le variabili dipendenti misurano il comportamento umano. Il termine dipendenti fa riferimento al fatto che dipende da cosa fa il partecipante. Alcuni esempi possono essere: il tempo di completamento di un task, velocità, accuratezza, error rate, ecc... Informalmente, una variabile dipendente non è altro che un comportamento misurabile o osservabile.

Data collection

I dati per le variabili dipendenti possono esser collezionati in diversi modi. Idealmente so posso utilizzare gli stessi software per effettuare gli esperimenti per tenere traccia dei timestamp, tempo di pressione di un tasto, bottone cliccato, ecc...

Abbiamo due tipi di logs:

- **High-level:** riporta direttamente i valori delle variabili dipendenti

```
2 [#]presented    transcribed presented_characters    transcribed_characters    input_time(sec) pause_time(sec) total_time(sec)
wpm msd numBksp numDelChars total_error cor_error    uncor_error
3 Mary had a little lamb Mary had a little lamb 22 22 15.228 0.0 15.267 14.184397163120567 0 2 8
0.2666666666666666 0.2666666666666666 0.0
4 suburbs are sprawling up everywhere suburbs are sprawling up everywhere 35 35 11.287 0.0 11.308 29.768760520953307 0
0 0 0.0 0.0 0.0
5 pay off mortgage for a house Pay off a mortgage for a house 30 30 9.106 0.0 9.123 35.58093564682627 1 0 0
0.0333333333333333 0.0 0.0333333333333333
6 taking the train is usually faster taking the train is usually faster 34 34 9.148 0.0 9.166 36.729339746392654 0
0 0 0.0 0.0 0.0
7 exceed the maximum speed limit exceed the maximum speed limit 30 30 12.513 0.0 12.54 23.01606329417406 0 1 5
0.14285714285714285 0.14285714285714285 0.0
8 the laser printer is jammed the laser printer is jammed 27 27 7.998 0.0 9.260 36.00900225056264 0 0 0 0.0 0.0 0.0
9 a big scratch on the tabletop a big scratch on the tabletop 29 29 8.49 0.0 8.507 39.57597173144876 0 0 0
0.0 0.0 0.0
10 Mary had a little lamb Mary a little lamb 22 18 13.056 0.0 13.075 13.786764705882353 4 2 7 0.3793103448275862
0.2413793103448276 0.13793103448275862
11 the music is better than it sounds the music is better than it sounds 34 34 9.57 0.0 9.599 38.87147335423197 0
0 0 0.0 0.0
12 microscopes make small things look big microscopes make small things look big 38 38 10.461 0.0 10.49
30.972182391740752 0 0 0 0.0 0.0 0.0
```

- **Low-level:** riporta tutti gli eventi salienti delle prove

```
1 #Log opened: Fri Apr 12 09:16:35 CEST 2013
2 #just like it says on the can good
3 #1365751074947
4 0,<Entr>,pos@0
5 220,<Entr>,pos@1
6 #REJ,
7 #just like it says on the can good
8 #1365751188459
9 0,j,pos@0
0 452,jus,pos@0
1 1219,just,pos@0
2 1970,just,pos@0
3 1981,<Sp>,pos@4
4 2561,l,pos@5
5 2828,li,pos@5
6 3567,lilo,pos@5
7 7975,lilo,pos@5
8 7988,<Entr>,pos@9
9 8451,<Entr>,pos@10
0 #REJ,just lilo
1 #just like it says on the can good
```

Entrambi i due tipi di log sono utili per l'analisi dei dati.

5.1.3 Control variable

Le variabili di controllo sono variabili mantenute costanti durante quei test in cui si cambiano le test condition delle variabili indipendenti al fine di capirne l'effetto sulle variabili dipendenti (tale test prende il nome di test-effect).

Esempio: Research question: Is there an effect of font color or background color on reading comprehension?

- **IV:** font color, background color;
- **DV:** comprensione dei test scores;
- **Control variables:** font size, font family, luce ambientale, ecc...

5.1.4 Random variable

Le variabili random sono delle circostanze che mutano casualmente. Possono essere associate al caso in cui si parla di una caratteristica dell'utente che è casuale ad esempio il nervosismo, il grado di stress ecc...

Esempio: Research question: Does user stance affect performance while playing Guitar Hero?

- **IV:** stance (standing, sitting);
- **DV:** score on songs;
- **Random variables:** esperienza pregressa come musicista, esperienza pregressa a giocare a guitar hero, numero di caffè consumati prima dei test, ecc...

Control vs Random Variables

Esiste un compromesso che può essere esaminato in termini di validità interna e validità esterna:

| Variabile | Vantaggio | Svantaggio |
|----------------|--|--|
| Random | Migliora la validità esterna utilizzando una varietà di situazioni e persone. | Compromette la validità interna introducendo ulteriore variabilità nei comportamenti misurati. |
| Control | Migliora la validità interna poiché la variabilità dovuta a una circostanza controllata viene eliminata. | Compromette la validità esterna limitando le risposte a situazioni e persone specifiche. |

5.1.5 Confounding variable

Le variabili confounding sono quelle variabili che al mutare di una variabile indipendente mutano anche loro.

Esempio: Research question: In an eye tracking application, is there an effect of "camera distance" on task completion time?

- **IV:** Camera distance (near, far):
 - Near camera (A) : cheap camera mounted on eye glasses;
 - Far camera (B) : expensive camera mounted above system display;
- **DV:** task completion time;
- Il problema di questo esperimento è che la "camera" è una variabile di confounding. La camera A per il setup "near", e la camera B for "far" setup. In più c'è anche la differenza di costo. Quindi siamo sicuri che otterremo dei risultati giusti?

Nei nostri esperimenti non vogliamo l'esistenza di confounding variable.

Within-subjects e between-subjects

- **Within-subjects:** Abbiamo la presenza di n test conditions, e ogni partecipante farà tutte le n test conditions. I vantaggi di tale tecnica sono l'utilizzo di pochi partecipanti per fare una buona valutazione, non vi è bisogno di bilanciare il gruppo: tutti testano tutto. Mentre gli svantaggi sono che ogni test deve esser eseguito da ogni partecipante quindi il procedimento è più lungo, l'esecuzione di vari test porta ad avere interferenza tra di loro.
- **Between subject:** Abbiamo n test condition, vengono creati n gruppi a cui è assegnata una e solo una test condition. Il vantaggio è non avere interferenze fra i vari test perché ogni gruppo esegue soltanto una test condition. Gli svantaggi sono la necessità di dover avere più partecipanti, all'interno dei gruppi bisogna avere molta variazione, e in più i gruppi devono essere bilanciati.

Per ogni IV bisogna assegnare sia within subject che between subject.

5.2 Order effect, counterbalancing

Nel within subject esiste un problema definito **order effect** cioè la possibile confusione provocata ai partecipanti a causa del fatto che devono eseguire tutti gli esperimenti.

Questo può essere evitato tramite la tecnica chiama **counter balancing** che consiste nel dividere i partecipanti in un gruppo e assegnare i vari test ai gruppi in ordine differente per non provocare a tutti la stessa confusione.

5.2.1 Latin Squares

Un problema è come scegliere quest'ordine. Una delle possibili soluzioni è utilizzando la tecnica Latin square che consiste nel creare una tabella nxn dove n è il numero dei test condition:

| 2 x 2 | | 3 x 3 | | | | | | | |
|-------|---|-------|---|---|---|---|---|---|--|
| A | B | A | B | | | | | | |
| B | A | B | C | | | | | | |
| | | | | | | | | | |
| A | B | C | A | | | | | | |
| B | C | A | B | | | | | | |
| C | A | B | C | | | | | | |
| | | | | | | | | | |
| 4 x 4 | | 5 x 5 | | | | | | | |
| A | B | C | D | A | B | C | D | E | |
| B | C | D | A | B | C | D | E | A | |
| C | D | A | B | C | D | E | A | B | |
| D | A | B | C | D | E | A | B | C | |
| | | | | | | | | | |
| A | B | C | D | E | A | B | C | D | |
| B | C | D | A | | B | C | D | E | |
| C | D | A | B | | C | D | E | A | |
| D | A | B | C | | D | E | A | B | |
| | | | | | | | | | |

C'è un problema però, abbiamo più volte stesse sequenze ma con ordini diversi e questo provoca confusione (nella 4x4 = AB, BA, AB, AB).

5.2.2 Balanced Latin Squares

Una possibile soluzione la otteniamo con le **balanced latin square**, in cui ogni condizione precede e segue ogni altra condizione un numero uguale di volte. Garantisce un adeguato bilanciamento di variabili e fattori di influenza, facendo sì che ogni elemento compare una volta sola in ogni riga e in ogni colonna. È possibile realizzare questo sistema soltanto con valori pari.

| 4 x 4 | | | | 6 x 6 | | | | | |
|-------|---|---|---|-------|---|---|---|---|---|
| A | B | D | C | A | B | F | C | E | D |
| B | C | A | D | B | C | A | D | F | E |
| C | D | B | A | C | D | B | E | A | F |
| D | A | C | B | D | E | C | F | B | A |
| | | | | E | F | D | A | C | B |
| | | | | F | A | E | B | D | C |

5.2.3 Test conditions with multiple factors

È un metodo di test in cui vengono considerati simultaneamente più fattori che possono influenzare il risultato finale della prova. In altre parole, invece di testare un singolo fattore alla volta, vengono testati contemporaneamente più fattori per determinare come essi interagiscono tra di loro e come influenzano il risultato finale. Di norma si sceglie un sotto-insieme di possibili test da testare perché testarli tutti porta ad avere come numero di test un "prodotto cartesiano":

- A: A₁, A₂, ... , A_n;
- B: B₁, B₂, B_m

Il numero possibili di test sarà quindi n x m.

Esempio

In questo esperimento si vuole determinare di quanto tempo differiscono 3 metodi di editing (A, B, C). Il task è il seguente: "Sostituire una parola di 5 lettere con un'altra"

Le condizioni sono assegnate con within-subjects:

- **Method A**: arrow keys, backspace, type;
- **Method B**: search and replace dialog;
- **Method C**: point and double click with the mouse, type

20 partecipanti sono stati reclutati, e sono stati divisi in 3 gruppi. Come metodo di somministrazione è stata scelta una matrice latina 3 x 3.

| Participant | Test Condition | | | Group | Mean | SD | | | |
|-------------|----------------|-------|-------|----------|------|------|--|--|--|
| | A | B | C | | | | | | |
| 1 | 12.98 | 16.91 | 12.19 | 1 ABC | 14.7 | 1.84 | | | |
| 2 | 14.84 | 16.03 | 14.01 | | | | | | |
| 3 | 16.74 | 15.15 | 15.19 | | | | | | |
| 4 | 16.59 | 14.43 | 11.12 | 2 BCA | 14.6 | 2.46 | | | |
| 5 | 18.37 | 13.16 | 10.72 | | | | | | |
| 6 | 15.17 | 13.09 | 12.83 | | | | | | |
| 7 | 14.68 | 17.66 | 15.26 | | | | | | |
| 8 | 16.01 | 17.04 | 11.14 | | | | | | |
| 9 | 14.83 | 12.89 | 14.37 | | | | | | |
| 10 | 14.37 | 13.98 | 12.91 | 3 CAB | 14.4 | 1.88 | | | |
| 11 | 14.40 | 19.12 | 11.59 | | | | | | |
| 12 | 13.70 | 16.17 | 14.31 | | | | | | |
| Mean | 15.2 | 15.5 | 13.0 | | | | | | |
| SD | 1.48 | 2.01 | 1.63 | | | | | | |

Group effect is small
 ∴ Counterbalancing worked!

Il differenza del group effect è minimo quindi il counterbalancig ha funzionato, se differiscono di troppo bisogna ri-testare tutto.

5.2.4 Other techniques

Esistono anche altre tecniche che possono essere utilizzate, ricordiamo che una matrice latina ha ordine $n!$. Le condizioni possono essere randomizzate. La randomizzazione è migliore se i compiti sono brevi e ripetuti spesso.

Legge di Fitts suggerisce che la progettazione di un sistema dovrebbe prevedere l'organizzazione degli elementi interattivi in modo da minimizzare la distanza tra gli elementi frequentemente utilizzati e massimizzare la dimensione degli elementi di destinazione. Inoltre, le dimensioni dei pulsanti dovrebbero essere scelte in modo da consentire all'utente di interagire con l'applicazione in modo efficiente e preciso.

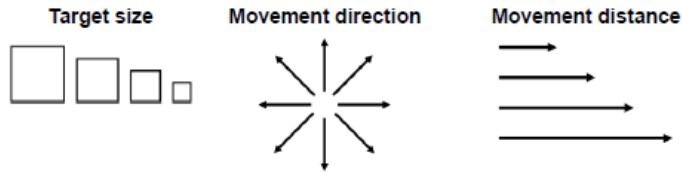


Figura 5.1: $4 \times 8 \times 4 = 128$ casi di test

5.3 Experiment Task

Quando si effettua un esperimento questo deve alla fine suscitare un cambiamento (elicit a change). Un buon task deve:

- **Represent:** attività che le persone possono fare con l'interfaccia;
- **Discrimante** (discriminare): tra le condizioni di prove;

5.3.1 Knowledge-based Task

La maggior parte dei task sperimentali sono basati sulle prestazioni o sulle abilità (ad esempio, inserire un'equazione, programmare una posizione di destinazione;). A volte l'attività è basata sulla conoscenza (ad esempio, "Usa un'interfaccia di ricerca su Internet per trovare la data di nascita di Albert Einstein."). In questo caso, i partecipanti si contaminano dopo la prima esecuzione del compito, poiché hanno acquisito la conoscenza. Per risolvere questo problema si può ripetere lo stesso task ma variandolo di poco (ad esempio, cercare la data di nascita di William Shakespeare")

Procedure Una procedura descrive tutto l'occorrente e i passi che devono esser fatti nei vari test. Essa può comprendere anche la descrizione dettagliata dei task da eseguire.

Istruzioni È molto importante dare istruzioni ai partecipanti se l'obiettivo del nostro esperimento è "procedere il più rapidamente e accuratamente possibile ma a un ritmo confortevole", devono essere date le stesse istruzioni a tutti i partecipanti. Se un partecipante chiede chiarimenti, non bisogna modificare le istruzioni in un modo che possa indurre il partecipante a comportarsi diversamente dagli altri.

Partecipanti I ricercatori vogliono che i risultati sperimentali si applichino a persone non effettivamente testate: una popolazione. Esempi di popolazione (Adulti, adolescenti, bambini, persone con determinate disabilità, mancini, ingegneri, musicisti, ecc.). Affinché i risultati si applichino generalmente a una popolazione, i partecipanti utilizzati nell'esperimento devono essere membri della popolazione desiderata e scelti a caso dalla popolazione. Il vero campionamento casuale viene eseguito raramente, tipicamente si utilizza una qualche forma di campionamento.

La scelta di quanti partecipanti è importante:

- **Troppo pochi (few):** gli effetti sperimentali non raggiungono la significatività statistica;
- **Troppi (many):** significatività statistica per effetti privi di valore pratico;
- **Numero corretto:** usa lo stesso numero di partecipanti utilizzato in una ricerca simile. Considera anche lo schema di controbilanciamento.

5.4 Questionari

Sono utilizzati in molti esperimenti dell'HCI. Hanno due scopi:

- Raccogliere informazioni sui partecipanti:

- Dati demografici (sesto, età, prima lingua, manualità, acuità visiva, ecc.);
- Precedente esperienza con interfacce o tecniche di interazione relative alla ricerca.

- Sollecitare feedback, commenti, impressioni, suggerimenti, ecc. sull'uso dell'apparato sperimentale da parte dei partecipanti.

I questionari se fatti devono essere brevi.

| | |
|--|--------------------|
| Which browser do you use? _____ | Open-ended |
| Which browser do you use? <input type="checkbox"/> Mozilla Firefox <input type="checkbox"/> Google Chrome <input type="checkbox"/> Microsoft IE <input type="checkbox"/> Other (_____) | Closed |
| Please indicate your age: _____ | Ratio-scale data |
| Please indicate your age? <input type="checkbox"/> < 20 <input type="checkbox"/> 20-29 <input type="checkbox"/> 30-39 <input type="checkbox"/> 40-49 <input type="checkbox"/> 50-59 <input type="checkbox"/> 60+ | Ordinal-scale data |

Using NASA Task Load Index (TLX):

| | | | | | | |
|---|---|---------|---|-------------------|---|---|
| Frustration: I felt a high level of insecurity, discouragement, irritation, stress, or annoyance. | | | | | | |
| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| Strongly disagree | | Neutral | | Strongly agree | | |

ISO 9241-9:

| | | | | | | |
|--------------|---|---|---|---|-------------|---|
| Eye fatigue: | | | | | | |
| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| Very high | | | | | Very low | |

5.5 System Usability Scale

Composto da 10 domande con punteggio fino a 5 o a 7(Likert scale). Le domande in posizione dispari hanno peso positivo, quelle in posizione pari hanno peso negativo. Il punteggio finale è in un range 0-100:

1. Sommare il contributo di ogni domanda (0 - 4), il contributo si calcola come segue:

- Per le domande 1,3,5,7 e 9 il contributo è dato dal punteggio dato dall'utente - 1 (questo per ogni domanda): es (3 -1, 2 - 1, 4 -1, ...), infine si sommano tra di loro;
- Per le domande 2, 4, 6, 8 e 10 il contributo è dato da 5 - il punteggio della domanda, alla fine si sommano tra di loro.

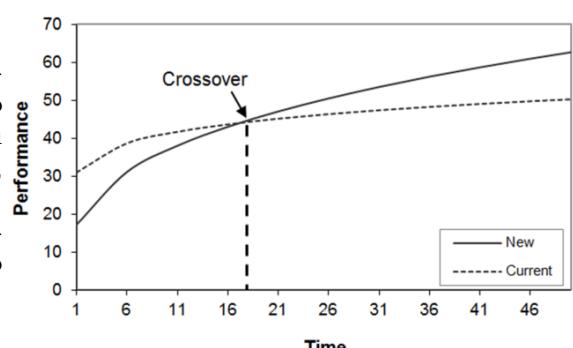
2. Infine si moltiplica la somma dei due punteggi ottenuti precedentemente con 2,5 e si ottiene così il valore del SU.

SUS è diventato uno standard del settore, con riferimenti in oltre 1300 articoli e pubblicazioni. È una scala molto semplice da somministrare ai partecipanti. Può essere utilizzato su campioni di piccole dimensioni con risultati affidabili. È valido: può distinguere efficacemente tra sistemi utilizzabili e inutilizzabili.

5.6 Longitudinal Studies

A volte, invece di "bilanciare" gli effetti dell'apprendimento, la ricerca cerca di promuovere e indagare l'apprendimento. In tal caso, viene condotto uno studio longitudinale. "La Pratica" è la IV. I partecipanti effettuano gli esperimenti per un periodo di tempo prolungato: Unità pratiche: blocchi, sessioni, ore, giorni, ecc.

A volte una nuova tecnica inizialmente funzionerà male rispetto a una tecnica consolidata. Uno studio longitudinale determinerà se si verifica un punto di **crossover** e, in tal caso, dopo quanta pratica.



5.7 Esperimento fatto in classe (Tapping VS Gesture Typing on Smartphones)

L'esperimento si basa sull'inserimento del testo (text entry). Quest'operazione è molto facile sui computer, ma è molto scomoda e inefficiente sui dispositivi mobili per via delle loro dimensioni piccole e per via della mancanza del feedback tattile (Lack of tactile feedback).

Le tastiere più utilizzate sono della tipologia QWERTY, altre soluzioni utilizzano layout OPTI (Figura 5.2), T9 per correggere eventuali errori, speech recognition o handwriting recognition.

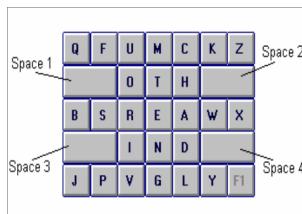


Figura 5.2: Figura 5.2

Esperimento

L'obiettivo dell'esperimento è determinare quale metodo di scrittura è più veloce/accurato/comodo fra Tapping e Gesturing.

Gesture writing : Il nome commerciale di questo metodo è Swipe. L'idea è quella di migliorare l'efficienza di scrittura effettuando degli shift invece che dei taping, portando ad un miglioramento con la pratica. È stato introdotto per la prima volta nel 2003/2004 da Christensson and Zhai.

Experiment design

- Partecipanti: tu;
- Dispositivo : Il tuo smartphone e un software che analizza le metriche durante l'immissione del testo.
- Design: within-subjects:
 - IV: Metodo { Tapping, Gesturing }, Sessioni { 1, 2, 3 };
 - DV : Velocità e Error Rate (per l'accuratezza).

5.7.1 Come calcolare la performance delle metriche nella text entry?

Speed

La **velocità** a volte è misurata in caratteri per secondo (cps): $cps = (\#chars / time)$.

Negli esperimenti però si utilizzano le **parole al minuto (wpm)**. Sappiamo che un minuto contiene 60 secondi, una parola convenzionalmente è di 5 caratteri, quindi $wpm = cps \times 60 / 5 = cps * 12$

Accuracy

Abbiamo due testi **A** e **B**. A è il testo che il partecipante deve immettere, B è quello che inserisce. Per vedere il numero di errori effettuati si utilizza il **MSD**.

MSD(P, T) è il minimo numero di operazioni per trasformare T in P, le operazioni possibili sono *inserimento*, *cancellazione* o *sostituzione di un singolo carattere*. Quindi l'**Error Rate** si calcola nel seguente modo:

$$\text{MSD Error Rate} = \frac{\text{MSD}(A, B)}{\text{MSD}(|A|, |B|)} * 100\% \quad (5.1)$$

NOTA: $\text{MSD}(|A|, |B|)$ si sceglie la parola con la lunghezza minima.

Total Error Rate (TER): tutti gli errori commessi dai partecipanti, sia quelli corretti che quelli non corretti.

Not Corrected Error Rate (NCER): errore ancora presente nella trascrizione del testo.

Hypothesis testing

L'obiettivo principale di questa tipologia di test è quello di analizzare i dati dei nostri esperimenti, utilizzando un semplice tool per capire se i dati ottenuti sono significanti o meno.

L'hypothesis testing non utilizza la stessa domanda che si pone la ricerca ma la modifica un po':

- Domanda generica di una ricerca:

Is the time to complete a task less using Method A than using method B?

- Nell'hypothesis testing la domanda della ricerca diventa uno statement:

There is no difference in the mean time to complete a task using method A vs Method B

Questo tipo di statement è chiamato **null hypothesis** (ed assume che non ci sia differenza).

Una procedura statistica cerca di rifiutare o accettare l'hypothesis. Quello che noi vogliamo è che sia rifiutata.

6.1 Statistical procedures

Esistono due tipologie di procedure statistiche:

- **Parametrica:** i dati provengono da una qualche distribuzione (Ratio e interval data);
- **Non-parametrica:** dati non provengono da nessuna distribuzione (Ordinal e nominal data, uso limitato per ratio e interval data).

| Measurement Scale | Defining Relations | Examples of Appropriate Statistics | Appropriate Statistical Tests |
|-------------------|---|--|--|
| Nominal | • Equivalence | • Mode • Frequency | • Non-parametric tests |
| Ordinal | • Equivalence • Order | • Median • Percentile | |
| Interval | • Equivalence • Order • Ratio of intervals | • Mean • Standard deviation | • Parametric tests • Non-parametric tests |
| Ratio | • Equivalence • Order • Ratio of intervals • Ratio of values | • Geometric mean • Coefficient of variation | |

Un esempio di queste procedure sono:

- **L'analisi della varianza (ANOVA):** utilizzato per ratio e interval data, è una delle procedure statistiche più utilizzate nella ricerca dell'HCI;
- **Post hoc tests:** utile nel caso in cui abbiamo più di 2 test conditions, utile anche per capire tra quale coppia del test c'è un effettiva differenza significativa.

Ci concentreremo su:

- **One-way ANOVA (single factor):** può avere 2 test conditions, > 2 test conditions (c'è bisogno del post-hoc test), 2 test conditions con Between-subject design;
- **Two-way ANOVA (two factors).**

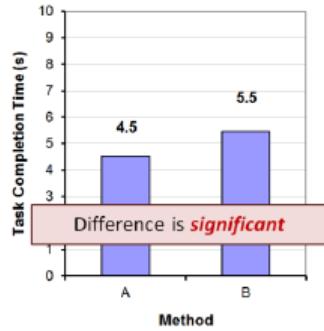
6.2 Analisi della varianza (ANOVA)

L'obiettivo di questo test è quello di determinare se una variabile indipendente ha un effetto significativo su una variabile dipendente. Il test ANOVA produce un valore chiamato **F-statistics**.

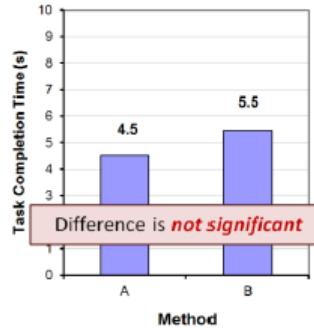
Mostreremo il funzionamento di questo test su due esempi che si basano sulla seguente domanda:

Is the time to complete a task less using Method A than using Method B?

Example #1



Example #2



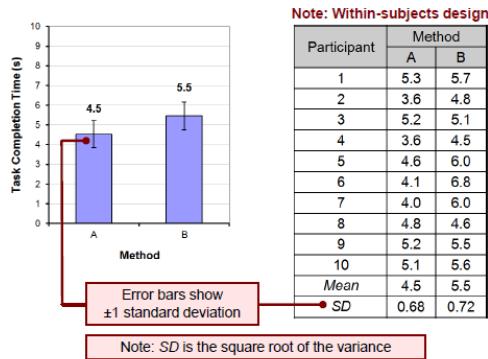
- **Significant**: implica che con ogni probabilità (likelihood) la differenza osservata è dovuta alle condizioni di prova;
- **Not significant**: implica che la differenza osservata è probabilmente dovuta al caso.

Ma come è possibile che due grafici identici, sullo stesso esperimento diano risultati diversi? Probabilmente è dovuto alla **deviazione standard (SD)**.

NOTA: la deviazione standard è una misura statistica che indica quanto i dati in un insieme si discostano dalla media di tale insieme.

In entrambi gli esempi abbiamo due variabili indipendenti che sono Il metodo A e il metodo B.

Esempio 1

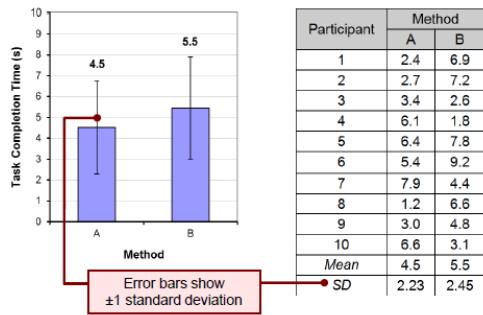


ANOVA Table for Task Completion Time (s)

| | DF | Sum of Squares | Mean Square | F-Value | P-Value | Lambda | Power |
|------------------|----|----------------|-------------|---------|---------|--------|-------|
| Subject | 9 | 5.080 | .564 | | | | |
| Method | 1 | 4.232 | 4.232 | 9.796 | .0121 | 9.796 | .804 |
| Method * Subject | 9 | 3.888 | .432 | | | | |

F-Value è il rapporto delle varianze, mentre **P-Value** rappresenta la probabilità di ottenere lo stesso risultato (o un risultato più estremo) se l'ipotesi nulla è vera. **DF** sono i gradi di libertà.

Nel primo esempio il valore di F viene riportato come $F_{1,9} = 9.80$, $p < 0.5$, dove le soglie per il valore di p sono: .05, .01, .005, .001, .0005, .0001.

Esempio 2**ANOVA Table for Task Completion Time (s)**

| | DF | Sum of Squares | Mean Square | F-Value | P-Value | Lambda | Power |
|------------------|----|----------------|-------------|---------|---------|--------|-------|
| Subject | 9 | 37.372 | 4.152 | | | | |
| Method | 1 | 4.324 | 4.324 | .626 | .4491 | .626 | .107 |
| Method * Subject | 9 | 62.140 | 6.904 | | | | |

Nel secondo esempio il valore di F viene riportato come $F_{1,9} = 0.626$, ns.

NOTA: per gli effetti non significanti si utilizza ns se $F < 1.0$ altrimenti si utilizza p > .05 se $F > 1.0$ e p > .05 .

6.2.1 Come si fa il resoconto (report) della F-statistic

- Deve essere messa tra parentesi;
- Maiuscolo (uppercase) per F e minuscolo (lowercase) per p;
- Usare il corsivo(italics) per F e p;
- Spazio in tutte e due lati dei segni = e <;
- Spazio dopo la virgola;
- I gradi di libertà devono essere scritti come pedici di F, plain font, carattere più piccolo;
- massimo 3 caratteri per il valore della F-statistic;
- Nessun 0 prima del punto per la p (tranne in Europa).

Di seguito un esempio di come fare il resoconto in un paper dei due valori precedentemente ottenuti:

- **Esempio 1:**

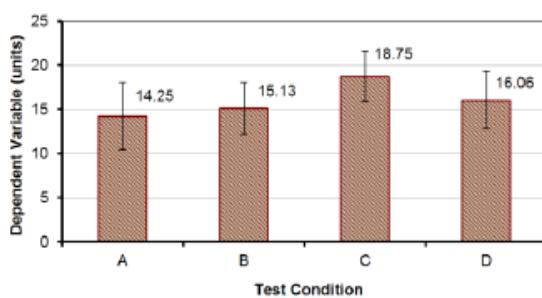
The mean task completion time for Method A was 4.5 s. This was 20.1% less than the mean of 5.5 s observed for Method B. The difference was statistically significant ($F_{1,9} = 9.80$, $p < .05$)

- **Esempio 2:**

The mean task completion time were 4.5 s for Method A and 5.5 s for Method B. As there was substantial variation in the observations across partecipants, the difference was not statistically significant as revealed in an analysis of variance ($F_{1,9} = 0.626$, ns)

6.3 Più di due test condition

| Participant | Test Condition | | | |
|-------------|----------------|-------|-------|-------|
| | A | B | C | D |
| 1 | 11 | 11 | 21 | 16 |
| 2 | 18 | 11 | 22 | 15 |
| 3 | 17 | 10 | 18 | 13 |
| 4 | 19 | 15 | 21 | 20 |
| 5 | 13 | 17 | 23 | 10 |
| 6 | 10 | 15 | 15 | 20 |
| 7 | 14 | 14 | 15 | 13 |
| 8 | 13 | 14 | 19 | 18 |
| 9 | 19 | 18 | 16 | 12 |
| 10 | 10 | 17 | 21 | 18 |
| 11 | 10 | 19 | 22 | 13 |
| 12 | 16 | 14 | 18 | 20 |
| 13 | 10 | 20 | 17 | 19 |
| 14 | 10 | 13 | 21 | 18 |
| 15 | 20 | 17 | 14 | 18 |
| 16 | 18 | 17 | 17 | 14 |
| Mean | 14.25 | 15.13 | 18.75 | 16.06 |
| SD | 3.84 | 2.94 | 2.89 | 3.23 |



Sulla test condition della variabile dipendente abbiamo un effetto molto significativo ($F_{3,45} = 4.95$, $p < .005$).

Gradi di libertà (DF), se n è il numero delle test conditions e m è il numero dei partecipanti, i gradi di libertà saranno:

- **Effect:** $(n-1)$, nel nostro caso $4-1 = 3$;
- **Residual:** $(n-1)(m-1) = 3 * 15 = 45$.
- Questo vale per un single-factor, in un design within-subjects.

ANOVA Table for Dependent Variable (units)

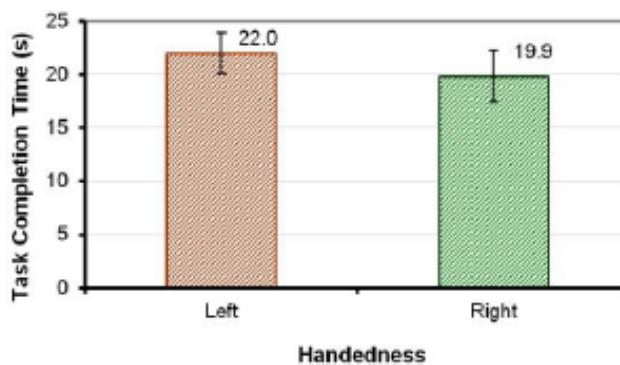
| | DF | Sum of Squares | Mean Square | F-Value | P-Value | Lambda | Power |
|--------------------------|----|----------------|-------------|---------|---------|--------|-------|
| Subject | 15 | 81.109 | 5.407 | | | | |
| Test Condition | 3 | 182.172 | 60.724 | 4.954 | .0047 | 14.862 | .896 |
| Test Condition * Subject | 45 | 551.578 | 12.257 | | | | |

6.3.1 Between-subject Designs

La domanda della ricerca è: Do left-handed users and right-handed users differ in the time to complete an interaction task?
La variabile indipendente è "handedness"

| Participant | Task Completion Time (s) | Handedness |
|-------------|--------------------------|------------|
| 1 | 23 | L |
| 2 | 19 | L |
| 3 | 22 | L |
| 4 | 21 | L |
| 5 | 23 | L |
| 6 | 20 | L |
| 7 | 25 | L |
| 8 | 23 | L |
| 9 | 17 | R |
| 10 | 19 | R |
| 11 | 16 | R |
| 12 | 21 | R |
| 13 | 23 | R |
| 14 | 20 | R |
| 15 | 22 | R |
| 16 | 21 | R |
| Mean | 20.9 | |
| SD | 2.38 | |

| Handedness | Task Completion Time (s) | |
|------------|--------------------------|------|
| | Mean | SD |
| Left | 22.0 | 1.93 |
| Right | 19.9 | 2.42 |



ANOVA Table for Task Completion Time (s)

| | DF | Sum of Squares | Mean Square | F-Value | P-Value | Lambda | Power |
|------------|----|----------------|-------------|---------|---------|--------|-------|
| Handedness | 1 | 18.063 | 18.063 | 3.781 | .0722 | 3.781 | .429 |
| Residual | 14 | 66.875 | 4.777 | | | | |

Non c'è una differenza statisticamente significante ($F_{1,14} = 3.78$, $p > .05$).

Gradi di libertà (DF), se n è il numero delle test conditions e m è il numero dei partecipanti, i gradi di libertà saranno:

- **Effect:** $(n-1)$, nel nostro caso $2-1 = 1$;
- **Residual:** $(m-n) (m-1) = 16 - 2 = 14$.
- Questo vale per un single-factor, in un design between-subjects.

6.4 Post Hoc Comparisons Tests

Un F-test significativo significa che almeno una delle condizioni del test differisce significativamente da un'altra condizione del test, ma questo non indica quale condizione di test differisce da quale. Per determinare quale coppia differisce significativamente abbiamo bisogno di utilizzare un test chiamato post hoc (Fisher PLSD, Bonferroni/Dunn, Dunnett, Tukey/Kramer, Games/Howell, Student-Newman-Keuls, orthogonal contrasts, Scheffé).

6.4.1 Scheffé Post Hoc comparisons

La comparazione di Scheffé è un tipo di analisi statistica utilizzata per confrontare le medie di più di due gruppi in uno studio. Questo tipo di analisi viene utilizzato quando l'analisi di varianza (ANOVA) indica una differenza significativa tra i gruppi, ma non fornisce informazioni specifiche su quali gruppi sono significativamente diversi.

| Scheffé for Dependent Variable (units) | | |
|--|------------|-------------|
| Effect: Test Condition | | |
| Significance Level: 5 % | | |
| | Mean Diff. | Crit. Diff. |
| A, B | -.875 | 3.302 |
| A, C | -4.500 | 3.302 |
| A, D | -1.813 | 3.302 |
| B, C | -3.625 | 3.302 |
| B, D | -.938 | 3.302 |
| C, D | 2.688 | 3.302 |

S

Test conditions A:C and B:C differiscono in modo significativo.

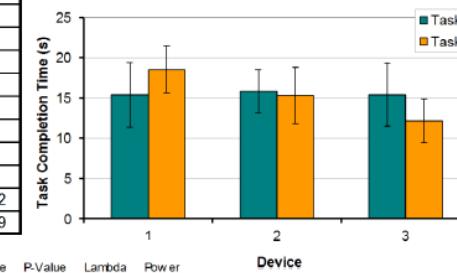
6.5 Two-way ANOVA

Un esperimento con due variabili indipendenti è chiamato two-way design. Viene effettuato un ANOVA test per i due effetti principali + l'unico effetto di interazione fra le due variabili. **Esempio:**

- Le variabili indipendenti sono i Device (D1,D2,D3) e Task(T1,T2);
- La variabile dipendente è il tempo di completamento del task;
- L'esperimento segue un design within-subjects;
- 12 partecipanti;

| Participant | Device 1 | | Device 2 | | Device 3 | |
|-------------|----------|--------|----------|--------|----------|--------|
| | Task 1 | Task 2 | Task 1 | Task 2 | Task 1 | Task 2 |
| 1 | 11 | 18 | 15 | 13 | 20 | 14 |
| 2 | 10 | 14 | 17 | 15 | 11 | 13 |
| 3 | 10 | 23 | 13 | 20 | 20 | 16 |
| 4 | 18 | 18 | 11 | 12 | 11 | 10 |
| 5 | 20 | 21 | 19 | 14 | 19 | 8 |
| 6 | 14 | 21 | 20 | 11 | 17 | 13 |
| 7 | 14 | 16 | 15 | 20 | 16 | 12 |
| 8 | 20 | 21 | 18 | 20 | 14 | 12 |
| 9 | 14 | 15 | 13 | 17 | 16 | 14 |
| 10 | 20 | 15 | 18 | 10 | 11 | 16 |
| 11 | 14 | 20 | 15 | 16 | 10 | 9 |
| 12 | 20 | 20 | 16 | 16 | 20 | 9 |
| <i>Mean</i> | 15.4 | 18.5 | 15.8 | 15.3 | 15.4 | 12.2 |
| <i>SD</i> | 4.01 | 2.94 | 2.69 | 3.50 | 3.92 | 2.69 |

| | Task 1 | Task 2 | Mean |
|----------|--------|--------|------|
| Device 1 | 15.4 | 18.5 | 17.0 |
| Device 2 | 15.8 | 15.3 | 15.6 |
| Device 3 | 15.4 | 12.2 | 13.8 |
| Mean | 15.6 | 15.3 | 15.4 |



| | ANOVA Table for Task Completion Time (s) | | | | | | |
|-------------------------|--|----------------|-------------|---------|---------|--------|-------|
| | DF | Sum of Squares | Mean Square | F-Value | P-Value | Lambda | Power |
| Subject | 11 | 134.778 | 12.253 | | | | |
| Device | 2 | 121.028 | 60.514 | 5.865 | .0091 | 11.731 | .831 |
| Device * Subject | 22 | 226.972 | 10.317 | | | | |
| Task | 1 | .889 | .889 | .076 | .7875 | .076 | .057 |
| Task * Subject | 11 | 128.111 | 11.646 | | | | |
| Device * Task | 2 | 121.028 | 60.514 | 5.435 | .0121 | 10.869 | .798 |
| Device * Task * Subject | 22 | 244.972 | 11.135 | | | | |

Il report sarà:

The grand mean for task completion time was 15.4 seconds. Device 3 was the fastest at 13.8 seconds, while device 1 was the slowest 17.0 seconds. The main effect device on task completion time was statistically significant ($(F_2, 22 = 5.865, p < .01)$. The task effect was modest, however. Task completion time was 15.6 seconds for task 1. Task 2 was slightly faster at 15.3 seconds; however, the difference was not statistically significant ($(F_1, 11 = 0.076, \text{ns})$. The results by device and task are shown in Figure x. There was a significant Device \times Task interaction effect ($(F_2, 22 = 3.78, p < .05$, which was due solely to the difference between device 1 task 2 and device 3 task 2, as determined by a Scheffé post hoc analysis.

Modeling Interaction

Un modello è una semplificazione di un oggetto reale.

Un modello può essere di diverso tipo, per esempio:

- **descrittivo:** fornisce informazioni sull'uso dello spazio, il movimento delle persone, la luce, l'ombra, ecc;
- **fisico:** è una previsione di trova si può trovare un oggetto in funzione del tempo.

7.1 Modelli predittivi

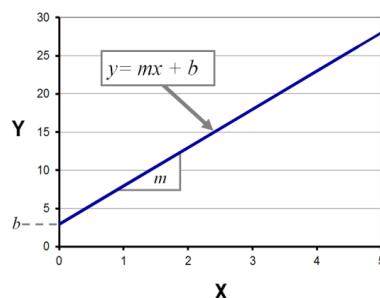
Un modello predittivo è un'equazione. L'equazione prevede il risultato di una variabile in base al valore di una o più altre variabili (predittori). Il risultato è una variabile dipendente, in genere il tempo o la velocità nell'esecuzione di un'attività.

La maggior parte delle fonti statistiche chiama la variabile preditrice la variabile indipendente.

I modelli predittivi sono utili nell' HCI, con un modello predittivo abbiamo a che fare numeri e non concetti.

7.1.1 Linear prediction equation

La più basica equazione di predizione è la relazione lineare tra la variabile preditrice x e una variabile di criterio y .



l'equazione di previsione lineare viene costruita utilizzando una procedura statistica nota come regressione lineare. Il suo obiettivo è dato un insieme di punti campione x-y, trovare i coefficienti m e b per la retta che minimizza le distanze al quadrato (minimi quadrati) dei punti dalla retta. Il risultato è un'equazione di previsione che fornisce la migliore stima di y in termini di x . Il presupposto, ovviamente, è che la relazione sia lineare.

Esempio

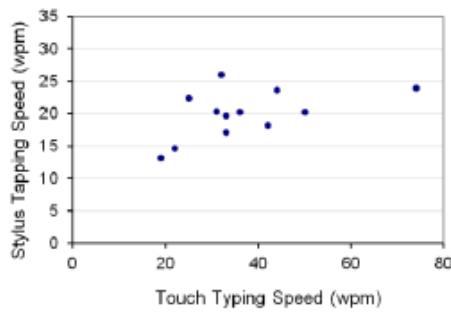
La domanda della ricerca è

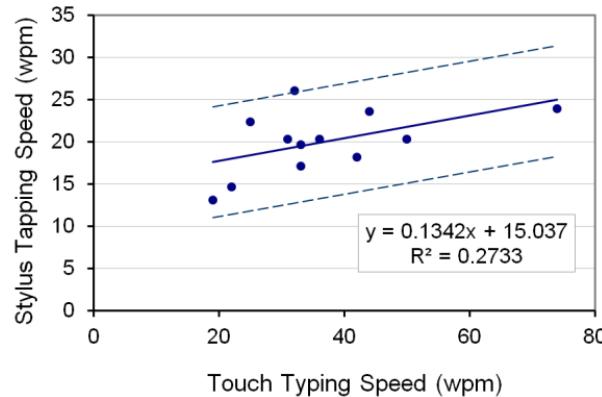
Can stylus tapping entry speed be predicted from touch typing entry speed?

La velocità di typing del touch è la variabile di predizione (x , misurata in un pre-test).

La velocità di typing dello stilo è la variabile di criterio (y , misurata sperimentalmente).

| Participant | Stylus Tapping Speed (wpm) | Touch Typing Speed (wpm) |
|-------------|----------------------------|--------------------------|
| P1 | 18.2 | 42 |
| P2 | 23.6 | 44 |
| P3 | 26.0 | 32 |
| P4 | 20.3 | 50 |
| P5 | 20.3 | 36 |
| P6 | 17.1 | 33 |
| P7 | 24.0 | 74 |
| P8 | 14.7 | 22 |
| P9 | 20.3 | 31 |
| P10 | 19.7 | 33 |
| P11 | 22.4 | 25 |
| P12 | 13.1 | 19 |



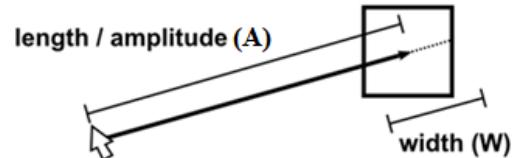


L'intervallo di confidenza è del 95% (è un intervallo di valori che rappresenta l'incertezza associata a una stima di un parametro statistico sulla base di un campione di dati). Nella foto sopra riportata troviamo anche l'equazione predittiva (y) e la correlazione quadratica (R^2). R^2 è un valore compreso tra 0 e 1 e indica la percentuale di variazione dei dati, nel nostro esempio 27% indica un valore negativo.

7.1.2 Fitt's law

Una delle più utilizzate nellHCI. È un modello utilizzato per i movimenti rapidi e mirati. La sua origine la si può trovare in due paper sperimentali nella psicologia, uno del 1954 e l'altro del 1964:

$$MT = a + b \log_2\left(\frac{A}{W} + 1\right) \quad (7.1)$$



I task possono essere classificati come serial task o discrete task a seconda della natura della selezione di oggetti sullo schermo.

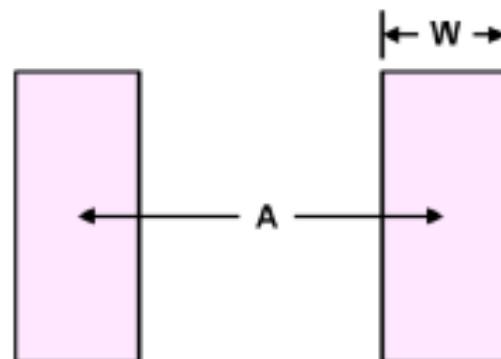
- Un **serial task** è un task in cui l'utente deve selezionare una serie di oggetti in sequenza, uno alla volta. In un serial task, la dimensione e la distanza dei singoli oggetti influenzano il tempo richiesto per selezionare l'intera sequenza di oggetti. Poiché ogni selezione dipende dalla selezione precedente, il tempo di selezione complessivo aumenta con il numero di oggetti da selezionare;
- Un **discrete task**, d'altra parte, è un task in cui l'utente deve selezionare un oggetto specifico tra un insieme di oggetti presenti sullo schermo. In un discrete task, la dimensione e la distanza dell'oggetto desiderato influenzano il tempo richiesto per selezionarlo. Poiché l'utente deve selezionare solo un oggetto alla volta, il tempo di selezione rimane costante indipendentemente dal numero di oggetti presenti.

In entrambi i casi, il tempo di selezione degli oggetti dipende dalla dimensione dell'oggetto e dalla distanza dal punto di partenza del cursore. La legge di Fitt afferma che il tempo di selezione è inversamente proporzionale alla distanza dell'oggetto e proporzionale al logaritmo della sua dimensione. In altre parole, oggetti più grandi e più vicini richiedono meno tempo per essere selezionati rispetto a oggetti più piccoli e più lontani.

Indice di fitt's della difficoltà (ID)

L'indice di difficoltà di Fitt's (ID) misura la difficoltà nel selezionare un target in un task.

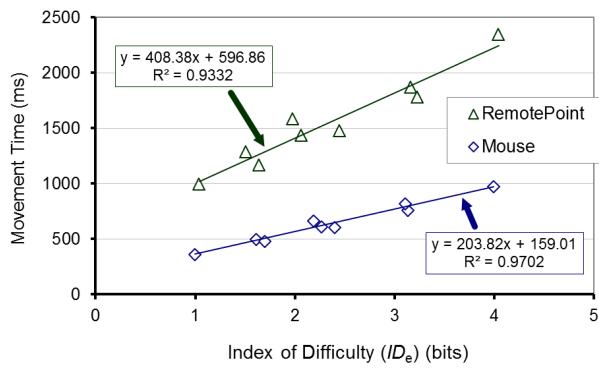
$$ID = \log_2\left(\frac{A}{W} + 1\right) \quad (7.2)$$



Fitts ha ipotizzato che la relazione tra il *movement time* (MT) e ID è lineare: $MT = a + b ID$

Esempio

Prendiamo in esempio un progetto che vuole comparare 2 dispositivi di puntamento. L'esperimento è stato fatto su 20 partecipanti.



7.1.3 Choice reaction time

Dati n stimoli, associati uno a uno con n risposte, il tempo di reazione all'inizio di uno stimolo è il tempo di reazione della scelta (choice reaction time).

Modellato da Hick-Hyman:

$$RT = a + \log_2(n + 1) \quad (7.3)$$

I coefficienti sono:

- $a \sim 200$ ms
- $b \sim 150$ ms

è un modello di elaborazione delle informazioni.

7.1.4 Keystroke-level model (KLM)

Uno dei primi e più completi modelli di HCI. Sviluppato specificamente per prevedere le prestazioni umane con sistemi informatici interattivi. Prevede i tempi di completamento delle attività senza errori da parte degli esperti.

Il KLM può prevedere il tempo per ciascun metodo.

Se utilizzato in fase di progettazione, possono essere prese in considerazione le alternative di progettazione e quindi seguire le scelte di progettazione confrontate.

Predizione di KLM

Innanzitutto un tast è suddiviso in una serie di subtask. Il tempo di predizione totale è la somma del tempo dei sottotast:

$$t_{\text{EXECUTE}} = t_K + t_P + t_H + t_D + t_M + t_R \quad (7.4)$$

Gli operatori sono:

- K -> keystriking, P -> pointing, H -> homing
- D -> drawing, M -> mental prep, R -> system response

Alcuni operatori vengono omessi o ripetuti, a seconda dell'attività (ad esempio, se sono necessarie n operazioni di battitura, t_K diventa $n \times t_K$)

| Operator | Description | Time (s) |
|-----------------|--|--|
| K | PRESS A KEY OR BUTTON Pressing a modifier key (e.g., shift) counts as a separate operation. Time varies with typing skill: Best typist (135 wpm) Good typist (90 wpm) Average skilled typist (55 wpm) Average non-secretary typist (40 wpm) Typing random letters Typing complex codes Worst typist (unfamiliar with keyboard) | 0.08 0.12 0.20 0.28 0.50 0.75 1.20 |
| P | POINT WITH A MOUSE Empirical value based on Fitts' law. Range from 0.8 to 1.5 seconds. Operator does <i>not</i> include the button click at the end of a pointing operation | 1.10 |
| H | HOME HAND(S) ON KEYBOARD OR OTHER DEVICE | 0.40 |
| D(n_D, l_D) | DRAW n_D STRAIGHT-LINE SEGMENTS OF TOTAL LENGTH l_D . Drawing with the mouse constrained to a grid. | .9 $n_D + .16 l_D$ |
| M | MENTALLY PREPARE | 1.35 |
| R(t) | RESPONSE BY SYSTEM Different commands require different response times. Counted only if the user must wait. | t |

Esempio

Consideriamo il seguente task:

Replace one 5-letter word with another, starting one line away

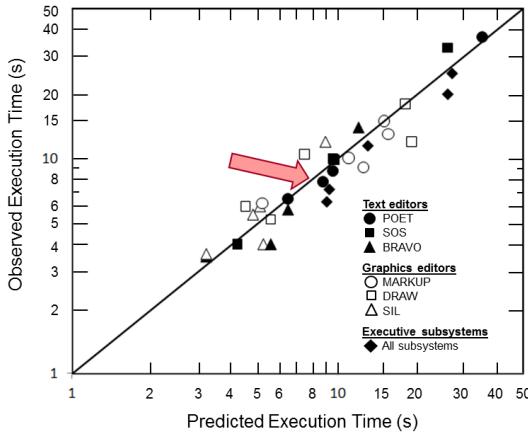
1. arrow keys, backspace, type;
2. search and replace dialog;
3. point and double click with the mouse, type.

| | |
|--------------------------|----------------------|
| Jump to next line | M K[LINEFEED] |
| Issue Substitute command | M K[S] |
| Type new word | K[word] |
| Terminate new word | M K[RETURN] |
| Type old word | K[word] |
| Terminate old word | M K[RETURN] |
| Terminate command | K[RETURN] |

$$t_{\text{EXECUTE}} = 4x t_M + 15x t_K \quad (7.5)$$

M la impostiamo come 1.35 secondi e K come 0.23 secondi.

$$t_{\text{EXECUTE}} = 4x 1.35 + 15x 0.23 = 8.85 \text{ secondi} \quad (7.6)$$



I valori sono quasi tutti sulla retta, quindi la predizione è giusta.

Applicazioni moderne

Quando è stato introdotto il KLM, l'interazione del mouse stava emergendo. Quindi oggi si può sostituire il valore costante del pointing (t_p), con il valore dell'equazione di predizione della Fitts'law:

$$t_p = 0.159 + 0.204x \log_2\left(\frac{A}{W} + 1\right) \quad (7.7)$$

Altro esempio

TASK: Change the font and style for "M K" to bold, Arial.

Con il mouse: Predizione:

| Mouse Subtasks | KLM Operators | t_P (s) |
|---|---------------|-----------|
| Drag across text to select "M K" | M P[2.5, 0.5] | 0.686 |
| Move pointer to Bold button and click | M P[13, 1] | 0.936 |
| Move pointer to Font drop-down button and click | M P[3.3, 1] | 0.588 |
| Move pointer down list to Arial and click | M P[2.2, 1] | 0.501 |
| $\sum t_P =$ | | 2.71 |

$$T_{EXECUTE} = 4xt_M + \sum t_P = 4x1.35 + 2.71 = 8.11 \text{ secondi.} \quad (7.8)$$

Con la tastiera: Predizione:

| Keyboard Subtasks | KLM Operators |
|--|--------------------|
| Select text | M K[shift] 3K[→] |
| Convert to boldface | M K[ctrl] K[b] |
| Activate Format menu and enter Font sub-menu | M K[alt] K[o] K[f] |
| Type a ("Arial" appears at top of list) | M K[a] |
| Select "Arial" | K[↓] K[enter] |

$$T_{EXECUTE} = 4xt_M + 12xt_K = 4x1.35 + 12x0.75 = 14.40 \text{ secondi.} \quad (7.9)$$

Per t_K è stato usato come tempo quello : "typing complex codes".

Esempio KLM e Mobile phone keypad

Con l'indice: $T_P = 0.165 + 0.052 \log_2\left(\frac{A}{W} + 1\right)$

Con il pollice: $T_P = 0.176 + 0.064 \log_2\left(\frac{A}{W} + 1\right)$

NOTA: non è possibile sapere precisamente quando un utente esegua un'operazione mentale (M), quindi ci sono due approcci da prendere includere una M operazione dove si ritiene giusto che possa essere eseguita oppure non includerla proprio. Questi due approcci producono un upperbound(akk-in) e lower-bound(all-out) nella predizione.

Research paper

La pubblicazione è la parte finale e essenziale di un progetto di ricerca. Un research paper si compone di diverse parti:

Title : Il titolo deve identificare l'oggetto del documento, restringere (narrow) l'ambito del lavoro e non deve essere ne troppo lungo.

Autore e affiliazioni :

| SIGCHI Conference Proceedings Format | | |
|--------------------------------------|-----------------------|-----------------------|
| 1st Author Name | 2nd Author Name | 3rd Author Name |
| Affiliation | Affiliation | Affiliation |
| Address | Address | Address |
| e-mail address | e-mail address | e-mail address |
| Optional phone number | Optional phone number | Optional phone number |

Abstract : Scritto per ultimo. Tipicamente contiene massimo 150 parole. È un singolo paragrafo senza citazioni. Deve descrivere brevemente cosa si è fatto e cosa si è trovato. Devono essere forniti i risultati più salienti.

Esempio:

What was done

This study addresses to what extent spatial mnemonics can be used to assist users to memorize or infer a set of text input chords. Users mentally visualize the appearance of each character as a 3x3 pixel grid. This grid is input as a sequence of three chords using one, two, or three fingers to construct each chord. Experiments show that users are able to use the strategy after a few minutes of instruction, and that some subjects enter text without help after three hours of practice. Further, the experiments show that text can be input at a mean rate of 5.9 words per minute (9.9 words per minute for the fastest subject) after 3 hours of practice. On the downside, the approach suffers from a relatively high error rate of about 10% as subjects often resort to trial and error when recalling character patterns.

(144 words)

What was found

Keywords : Utilizzate per favorire l'indicizzazione all'interno di un database e per la ricerca. Sono scelte dall'autore.

Computing Classification System Dal 1998, i documenti di conferenze e riviste ACM devono includere anche categorie, descrittori del soggetto e termini generali.

Introduction : È la prima sezione di un paper di ricerca. I contenuti che ci si aspetta sono:

- Introduzione al tema della ricerca;
- Problema dell'interfaccia utente o dichiarazione di sfida;
- Citazione delle soluzioni più importanti (se esistenti).
- Contributo del lavoro: Cosa c'è di nuovo e interessante nella ricerca? Anticipazione della soluzione imminente (che viene sviluppata e valutata nel resto del documento);
- Overview of paper.

Overview of Paper : Di solito alla fine dell'introduzione viene fornita una panoramica dell'intero documento.

Related Work : Revisione della letteratura: discussione del lavoro correlato, si descriva brevemente ogni lavoro, evidenzia le differenze e le somiglianze con la tua soluzione. Includere citazioni (con informazioni bibliografiche complete nella sezione di riferimento alla fine), possibilmente includere una tabella che riassume le caratteristiche principali del lavoro correlato. Classificare gli approcci correlati in base a una caratteristica importante comune, di solito, nei lavori brevi la sezione non è presente e un piccolo numero di lavori correlati è riassunto nell'introduzione.

Proposal : descrive i dettagli della soluzione proposta. Una tecnica interattiva è difficile da descrivere, si possono usare immagini, o riferimenti video (es: youtube).

Sezioni e sub-sezioni, non ci sono regole organizzate come vuoi.

Evaluation : Racconta al lettore come l'esperimento è stato progettato e realizzato. I titoli variano (valutazione, metodo, metodologia, esperimento, studio dell'utente, ...). La sezione di valutazione deve essere diretta: semplice, chiara, prevedibile (come una ricetta). La ricerca deve essere replicabile (come già notato). La sezione deve fornire informazioni sufficienti affinché un ricercatore esperto possa replicare l'esperimento se lo desidera.

Predictability : L'organizzazione della sezione di valutazione deve essere prevedibile. Questo consente a un lettore di setacciare rapidamente i documenti per trovare i punti chiave nella progettazione dell'esperimento. La convenzione impone che la sezione del metodo contenga le seguenti sottosezioni (e nel seguente ordine):

- Partecipanti: indica al lettore il numero di partecipanti e come sono stati selezionati: erano volontari o erano pagati? Vengono fornite anche informazioni demografiche (ad es. età, sesso, esperienza correlata, ...). Altri dettagli, se del caso (ad esempio, reddito, livello di istruzione più alto, acuità visiva, ...). Questa sezione è generalmente breve, tuttavia;
- Apparato: descrive il sistema (hardware e software). Le intestazioni variano (ad esempio, Materiali, Interfaccia, ...). La riproducibilità è estremamente importante, quindi bisogna fornire tutti i dettagli necessari. Usa schermate o foto dell'interfaccia. Se i dettagli tecnici sono stati divulgati nell'introduzione, è sufficiente rimandare il lettore a una sezione precedente (ad esempio, "il software includeva l'algoritmo descritto nella sezione precedente");
- Procedura: dice al lettore esattamente cosa è successo con ciascun partecipante. Cose da notare: istruzioni, descrizione del compito, dimostrazione o pratica, somministrazione del questionario, ripetizioni di prova, pause di riposo, tempo totale, ecc.

Questa sezione serve anche per descrivere l'**experiment task**.

Un'altra sezione è quella del design, che riassume l'esperimento in termini di variabili, assegnazione delle condizioni, ecc.

- Progetto.

Results and discussion : I risultati e la discussione sono solitamente combinati. Se sono presenti valori anomali o qualsiasi filtraggio o trasformazione dei dati, bisogna dichiararlo in anticipo. L'Approccio statistico e i test a volte sono convogliati in un paragrafo iniziale. Nessuna regola rigida, ma un approccio comune consiste nell'organizzare questa sezione in base a variabili dipendenti, iniziando dalle più importanti (ad es. velocità, tempo di completamento dell'attività). Per ogni variabile dipendente, inizia con un'osservazione ampia, quindi passa a dettagli più fini. Riporta i risultati dell'analisi statistica. Discuti e spiega i risultati ottenuti.

Non bisogna dare troppi risultati, compara i risultati con quelli degli altri lavori, usa gli schemi appropriati per rappresentare i dati, inserisci i feedback dei partecipanti.

Conclusion : Riassumi quello che hai fatto, riformula il contributo e/o i risultati significativi, identifica gli argomenti per ulteriori lavori (ma evitare di sviluppare nuove idee nella sezione Conclusione).

References : Informazioni bibliografiche complete per gli articoli citati. Formatta come richiesto (i dettagli contano!)

REFERENCES

1. Aula, A., Khan, R. M., and Guan, Z., How does search behavior change as search becomes more difficult? *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems - CHI 2010*, (New York: ACM, 2010), 35-44.
2. Brajnik, G., Yesilada, Y., and Harper, S., The expertise effect of web accessibility evaluation methods, *Human-Computer Interaction*, 26, 2011, 246-283.
3. Brown, T., *Change by design: How design thinking transforms organizations and inspires innovation*. New York: HarperCollins, 2009.
4. Buxton, W., There's more to interaction than meets the eye: Some issues in manual input, in *User centered system design: New perspectives on human-computer interaction*, (D. A. Norman and S. W. Draper, Eds.). Hillsdale, NJ: Erlbaum, 1986, 319-337.
5. ESA, Electronic Software Association, *Industry facts*, <http://www.thesa.com/facts/>, (accessed February 4, 2012).

Conference paper

Journal paper

Book

Book chapter

Internet document