

UNIVERSITÀ DEGLI STUDI DI TRIESTE

Dottorato di Ricerca in Psicologia Sperimentale

VIII ciclo

***Valutare i sistemi flessibili:
un approccio globale alla HCI***

Antonella De Angeli '66

CW

Walter Gerbino

Relatore:

Prof. Walter Gerbino

Walter Gerbino

Coordinatore del dottorato:

Prof. Walter Gerbino

Dipartimento di Psicologia

Università di Trieste

A mio padre e ai suoi computer

Ringraziamenti

Le idee e gli esperimenti descritti in questa tesi sono frutto di vari lavori interdisciplinari svolti in collaborazione con centri di ricerca nazionali e internazionali.

Innanzitutto vorrei dedicare un ringraziamento particolare a INSIEL S.p.A. e a tutte le persone che hanno voluto e sostenuto il *Laboratorio Tecnologie Cognitive* dove ho svolto la maggior parte del mio lavoro. Fra loro ho piacere di ricordare Walter Gerbino e Danilo Fum che in questi anni hanno supervisionato il mio lavoro con competenza e fiducia.

Particolarmente importanti sono stati anche i frequenti rapporti con il gruppo di Linguaggio Naturale dell'IRST di Trento e di Ergonomia del CRF nonché i preziosi insegnamenti ricevuti presso lo IPO di Eindhoven e il CHCC dell'Oregon Graduate Institute di Portland. A tutte le persone con cui in questi anni ho avuto il piacere di collaborare vanno i miei più sentiti ringraziamenti e la più sincera stima.

Vorrei dedicare però un ringraziamento particolare a Daniela Petrelli, l'*alter ego* informatico degli esperimenti di simulazione e di molte riflessioni sulla natura della HCI esposte in questo lavoro.

Grazie di cuore anche alla mia famiglia *informatica*, agli amici e in modo particolare a Ubaldo il cui contributo e sostegno è stato assolutamente impagabile.

L'uomo è misura di tutte le cose
(Protagora, *Sulla Verità*, fr. 1, Abbagnano, 1982)

INDICE

	pag.
Premessa.....	8
1. UTENTI E COMPUTER.....	12
1.1. Soluzioni di avvicinamento.....	13
1.2. Verso il dialogo flessibile.....	14
1.2.1. <i>Lo sviluppo delle interfacce.....</i>	<i>16</i>
1.3. I sistemi multimodali.....	19
1.3.1. <i>Interfacce multimodali.....</i>	<i>20</i>
1.4. Artefatti socio-cognitivi.....	21
1.5. Difficoltà di analisi dell'interazione flessibile.....	23
2. HCI: UNA DISCIPLINA DINAMICA.....	25
2.1. Human-Computer Interaction.....	26
2.1.1. <i>I motivi della scelta.....</i>	<i>28</i>
2.2. Oggetto di studio.....	30
2.2.1. <i>Alcune proposte.....</i>	<i>31</i>
2.3. L'approccio interdisciplinare.....	33
2.3.1. <i>Il contributo della psicologia.....</i>	<i>34</i>
2.4. Dal cognitivismo ai giorni nostri.....	36
2.4.1. <i>La psicologia ingegneristica.....</i>	<i>38</i>
2.4.2. <i>Oltre la psicologia ingegneristica.....</i>	<i>39</i>
2.5. Conclusioni.....	42
3. MISURE E METODOLOGIE DI VALUTAZIONE.....	43
3.1. Usabilità dei sistemi.....	44
3.1.1. <i>La proposta di Nielsen.....</i>	<i>45</i>
3.1.2. <i>La proposta ISO/CD 9241.....</i>	<i>46</i>
3.1.3. <i>Usability Engineering.....</i>	<i>47</i>
3.2. Valutazione basata sulla teoria.....	49
3.2.1. <i>Modelli teorici di massimizzazione dell'usabilità.....</i>	<i>50</i>
3.3. Valutazione basata sull'analisi del compito.....	51
3.3.1. <i>Metodi di tipo cognitivo.....</i>	<i>51</i>
3.4. Valutazione con osservatori reali.....	54
3.4.1. <i>Carenze metodologiche.....</i>	<i>56</i>
3.5. La simulazione come strumento conoscitivo.....	58
3.5.1. <i>Simulazione Wizard of OZ.....</i>	<i>59</i>
4. L'APPROCCIO GLOBALE ALLA HCI.....	62
4.1. La nuova proposta.....	63
4.1.1. <i>Pragmatica del metodo.....</i>	<i>64</i>
4.2. Caratterizzazione delle entità.....	65
4.2.1. <i>Ambiente.....</i>	<i>66</i>
4.2.2. <i>Compito.....</i>	<i>68</i>
4.2.3. <i>Computer.....</i>	<i>68</i>
4.3. Utente.....	70
4.3.1. <i>Un nuovo modo di concepire l'utente.....</i>	<i>71</i>
4.4. Analisi dei risultati.....	75
4.5. Conclusioni.....	77
5. LE DETERMINANTI RELATIVE ALL'UTENTE.....	79
5.1. Studio 1.....	81

5.1.1. Metodo.....	81
5.1.1.1. Popolazione della ricerca	81
5.1.1.2. Strumenti	81
5.1.1.3. Procedura.....	84
5.1.2. Risultati	84
5.1.2.1. Taratura degli strumenti.....	84
5.1.2.2. Verifica dell'ipotesi situazionalista	87
5.1.2.3. Predittori dell'esperienza	91
5.1.3. Conclusioni.....	92
5.2. Studio 2.....	93
5.2.1. Metodo.....	93
5.2.1.1. Popolazione della ricerca	93
5.2.2. Risultati	94
5.2.2.1. Verifica degli strumenti.....	94
5.2.2.2. Analisi del campione di scuola media inferiore	96
5.2.2.3. Predittori dell'esperienza	99
5.2.2.4. Paragone fra i due campioni.....	100
5.2.3. Conclusioni.....	100
5.3. Discussione.....	102
6. SIMULAZIONE 1.....	103
6.1. Definizione degli obiettivi.....	105
6.2. Caratterizzazione delle quattro entità.....	107
6.2.1. Caratterizzazione del computer.....	108
6.2.1.1. Modello interattivo	109
6.2.2. Caratterizzazione dell'utente	110
6.3. Selezione degli attributi di usabilità.....	112
6.4. Metodo.....	113
6.4.1. Partecipanti e disegno	113
6.4.2. Procedura	114
6.5. Risultati	115
6.5.1. Costruzione degli indici soggettivi.....	115
6.5.2. Omogeneità del campione.....	116
6.5.3. Ipotesi del vantaggio.....	117
6.5.4. Ipotesi di produzione	118
6.5.5. Facilità di apprendimento	119
6.5.6. Efficienza	123
6.6. Conclusioni.....	125
7. SIMULAZIONE 2.....	127
7.1. Definizione degli obiettivi.....	127
7.2. Caratterizzazione delle entità.....	128
7.3. Metodo.....	129
7.3.1. Partecipanti e disegno	129
7.3.2. Procedura.....	129
7.4. Risultati	131
7.4.1. Verifica delle ipotesi sperimentali.....	131
7.5. Conclusioni.....	133
7.6. Discussione finale	134
Riferimenti bibliografici.....	136

INDICE DELLE FIGURE

	pag.
<i>Figura 1. Ciclo di elaborazione dell'informazione secondo l'approccio cognitivista.</i>	37
<i>Figura 2. Modello dell'Elaboratore Umano (Adattato da Card et al., 1983, pag. 26). ...</i>	39
<i>Figura 3. Esempio di progettazione secondo i criteri stabiliti dalla Usability Engineering.</i>	48
<i>Figura 4. Determinanti dell'interazione secondo l'approccio globale.</i>	66
<i>Figura 5. Modello dell'attività umana (adattato da Norman, 1986 pag. 42).</i>	72
<i>Figura 6. Utilizzo di applicazioni informatiche in funzione del Sesso Biologico.</i>	89
<i>Figura 7. GDE e atteggiamento in funzione di Sesso Biologico e Identità Sessuale.</i>	90
<i>Figura 8. Giudizi di rassicurazione e amichevolezza in funzione di Uso e Sesso biologico.</i>	90
<i>Figura 9. Utilizzo di applicazioni informatiche in funzione del Sesso Biologico.</i>	97
<i>Figura 10. Utilizzo di applicazioni informatiche in funzione dell'Età.</i>	98
<i>Figura 11. GDE e Atteggiamento in funzione dell'Età.</i>	99
<i>Figura 12. Interfaccia dell'utente nella condizione IC Parziale e Feedback Presente. ...</i>	106
<i>Figura 13. Interfaccia del mago nelle condizioni IC Parziali.</i>	108
<i>Figura 14. Abilità I/O, atteggiamento e programmazione in funzione di IC.</i>	116
<i>Figura 15. Percentuali delle Strategie di Identificazione del Referente nelle due condizioni di IC.</i>	117
<i>Figura 16. Valori medi degli indici di discrepanza in funzione di Posizione Seriale e Ordine.</i>	120
<i>Figura 17. Numero totale di frasi in funzione di IC e Sesso.</i>	122
<i>Figura 18. Numero medio di frasi nei due compiti in funzione di IC e Feedback.</i>	123
<i>Figura 19. Numero medio di domande in funzione di IC e Feedback.</i>	124
<i>Figura 20. Percentuale di errori comunicativi nelle condizioni sperimentali.</i>	131
<i>Figura 21. Lunghezza delle frasi in funzione di Modulo e Sistema.</i>	132

INDICE DELLE TABELLE

	pag.
<i>Tabella 1. Principali differenze fra ipotesi disposizionalista e situazionalista.</i>	80
<i>Tabella 2. Taratura del CATT (20 item).</i>	86
<i>Tabella 3. Descrittive dei tre indici di atteggiamento verso il computer.</i>	87
<i>Tabella 4. Disponibilità del computer in funzione del sesso biologico.</i>	88
<i>Tabella 5. Regressione multipla di amichevolezza, rassicurazione e sesso su GDE.</i>	91
<i>Tabella 6. Distribuzione delle categorie di Identità Sessuale nei due campioni.</i>	96
<i>Tabella 7. Risultati analisi regressione.</i>	100
<i>Tabella 8. Valori medi di Paragone in funzione di Ordine.</i>	121

Premessa

A scientific psychology should not only help us to understand our own human nature, it should help us in our practical affairs.

(Card, Moran e Newell, 1983. *The Psychology of Human-Computer Interaction*, pag. 1).

Nella società in cui viviamo il computer si va imponendo come strumento di lavoro e oggetto di intrattenimento. La sua crescente importanza nella vita quotidiana lo ha reso oggetto di studio anche delle scienze umane, che si sono così trovate a contatto con discipline tecniche e ingegneristiche. Dal loro incontro è nata, all'inizio degli anni '80, la *Human-Computer Interaction* (da ora in poi, HCI¹), un campo di studio finalizzato alla comprensione e al miglioramento del sistema utente-computer, nel cui ambito si situa il mio lavoro di tesi.

In generale la HCI può essere collocata all'interno di un processo di interazione fra settori di studio diversi, che fin dall'immediato dopo guerra ha visto nascere nuove aree di ricerca. La psicologia, in quanto scienza dell'essere umano, è stata spesso attratta da queste prospettive di studio. Si pensi, per esempio, al suo ruolo nell'Intelligenza Artificiale (Bara, 1984; Job e Tabossi, 1984), nell'Ergonomia (Re, 1995), nelle Scienze dell'Informazione (Ingwersen, 1992) o, più recentemente, nelle Scienze Cognitive (Tabossi, 1988) e nel Connessionismo (Parisi, 1989). La tendenza allo scambio

¹ La scelta di denominare la disciplina con l'espressione inglese appare opportuna per vari motivi. Innanzitutto, perchè la HCI si è sviluppata prevalentemente in nazioni anglofone e tuttora in Italia non esiste una tradizione di ricerca consolidata e autonoma. Inoltre, la traduzione *Interazione Uomo-Computer* è poco soddisfacente sia per la sua valenza sessista sia per la mancanza di un acronimo con cui riassumere la lunga espressione.

interdisciplinare sembra frutto della consapevolezza che la psicologia da sola non è sufficiente a comprendere il comportamento umano o, per lo meno, a comprendere la complessità del comportamento umano fuori dal laboratorio.

Il desiderio di utilizzare specifiche conoscenze psicologiche nella vita quotidiana non è prerogativa esclusiva di domini applicativi, quali la HCI. Già Neisser nel suo *Cognition and Reality* (1976), auspicava uno spostamento dell'attenzione dalla ricerca di laboratorio al mondo esterno, alla vita reale delle persone. Lo stesso autore che nove anni prima aveva praticamente segnato l'inizio ufficiale del cognitivismo (Neisser, 1967), criticava il progressivo ripiegarsi della ricerca su se stessa e invitava a situare il concetto di informazione nell'ambiente in cui viene processata. L'interesse psicologico verso la HCI si basa su queste idee e le applica in un contesto che sempre maggiore importanza ha nella vita dell'essere umano.

In questa tesi mi propongo di dimostrare la necessità dello sviluppo di una specifica *psicologia della HCI* che superi le tradizionali barriere fra le varie aree di studio del comportamento umano, dando luogo a una *psicologia di confine* in cui confluisca parte dell'apparato teorico e metodologico della psicologia dei processi cognitivi, dell'ergonomia nonché della psicologia sociale. Sicuramente questo è un obiettivo molto ambizioso e problematico, ma come si cercherà di dimostrare nel corso dell'esposizione, sembra essere la soluzione più adatta per comprendere l'interazione fra gli utenti e i sistemi informatici di un futuro che si prospetta abbastanza prossimo.

Postulato fondamentale del lavoro è la natura *dinamica* della HCI. Il suo oggetto di studio tende, infatti, a modificarsi parallelamente allo sviluppo tecnologico lungo il continuum *vincolato-flessibile*. Ogni generazione di computer è caratterizzata da una maggiore flessibilità dialogica rispetto a quella precedente. In un dialogo vincolato l'utente deve adattarsi al sistema; in un dialogo flessibile è il sistema che si adatta all'utente, proponendo uno scambio più consono alle caratteristiche e alle abilità naturali dell'essere umano. L'evoluzione si riflette direttamente sul tipo di capacità e conoscenze richieste all'utente: il dialogo vincolato richiede quasi esclusivamente il coinvolgimento di abilità cognitive; quello flessibile anche di abilità sociali e comunicative. L'analisi dei sistemi interattivi flessibili necessita dunque un approccio e una metodologia più complessa di quella sufficiente per l'analisi dei sistemi vincolati.

L'*approccio globale*, sviluppato nelle pagine di questo lavoro, si propone come un contributo alla valutazione dei nuovi sistemi interattivi. L'idea innovativa della proposta è la definizione di *interazione utente-computer* come una *gestalt*, totalità dinamica prodotta dall'azione sinergica di quattro entità: utente, computer, compito e ambiente. In questa prospettiva, lo studio del processo interattivo non può essere ridotto, come tradizionalmente avviene, all'analisi separata delle entità che lo determinano, ma necessita la comprensione delle regole secondo cui tali entità danno luogo alla totalità. L'approccio

richiede dunque di combinare alla metodologia ergonomica, adeguata per lo studio dei processi cognitivi sottostanti l'interazione utente-sistema vincolato, strumenti della ricerca psicosociale, più idonei a rappresentare e comprendere la complessità dell'interazione essere umano-sistema flessibile.

La tesi si propone di presentare l'approccio globale da un punto di vista teorico e dimostrarne l'adeguatezza tramite due studi correlazionali e due esperimenti di simulazione volti alla valutazione di prototipi di sistemi multimodali. Il lavoro articola in due parti: una teorica, *Verso un nuovo modello*, che comprende i primi quattro capitoli, e una sperimentale *Applicazioni del modello*, che comprende gli ultimi tre.

Il primo capitolo, di carattere introduttivo, fornisce un quadro generale dello sviluppo tecnologico e delle sue conseguenze sull'interazione utente-computer. Il secondo presenta invece la HCI sottolineandone da vari punti di vista la natura dinamica e dedicando particolare attenzione al contributo della psicologia. Il capitolo successivo illustra le principali misure e metodologie con cui tradizionalmente è stata valutata la prestazione del sistema utente-computer. La parte teorica si chiude con la presentazione dell'approccio globale.

La parte sperimentale presenta due studi correlazionali, finalizzati a comprendere i reciproci rapporti fra alcune caratteristiche dell'utente che l'approccio ritiene di fondamentale importanza nell'interazione. I capitoli 6 e 7 descrivono invece due valutazioni di sistemi multimodali eseguite tramite la metodologia proposta dall'approccio globale.

Parte I
Verso un nuovo modello

1. Utenti e computer

Nella nostra vita quotidiana veniamo spesso a contatto con macchine più o meno complesse. Ci svegliamo al suono di radio sveglie digitali, ascoltiamo le notizie del giorno da sofisticati stereo, guidiamo potenti automobili, cuciniamo in forni a microonde, e infine ci rilassiamo guardando un bel film con il videoregistratore. Nel caso, poi, in cui svolgiamo un lavoro di ufficio, solitamente passiamo più tempo interagendo con delle macchine (solo per ricordare le più comuni, computer, fotocopiatrici, fax, telefoni cellulari e non) piuttosto che con i nostri colleghi.

Se qualcuno si sente contrariato da questa realtà, si rassegni. Il futuro sembra destinato a una crescente automazione delle attività umane. Nei laboratori di tutto il mondo sono in fase di studio nuovi sistemi che permetteranno di potenziare le attuali tecnologie informatiche e di applicarle anche in contesti diversi da quelli attuali (Nielsen 1993; Cole *et al.*, 1995). Solo per fare un esempio, computer di vario tipo entreranno presto a far parte della normale strumentazione delle automobili. Guideremo allora macchine che, oltre a permetterci spostamenti rapidi e sicuri, saranno capaci di fornirci informazioni sempre più affidabili e tempestive sullo stato del traffico, del tempo o sul percorso ottimale per raggiungere la meta desiderata (Bubb, De Angeli, Ferrante, Haller, Jacomussi, Toffetti e Varalda, 1994; Varalda, Bray, Morra, Sidoti, Gerbino, De Angeli e Ferrante, 1994; De Angeli, Fum e Gerbino, 1995; De Angeli, Fum, Gerbino, Pani, Parisatto e Tognoni, 1996).

L'inserimento pervasivo della tecnologia informatica nella vita di ogni giorno sta procedendo in due direzioni. Da un lato si cerca di *nascondere* il computer all'interno di oggetti d'uso quotidiano, per esempio notes e lavagne elettroniche nonché di progettare dispositivi meno intrusivi degli attuali, quali telecamere che seguono i movimenti dell'utente per selezionare automaticamente l'inquadratura ottimale². Dall'altro ci si orienta invece verso la costruzione di sofisticate interfacce, capaci di rendere le situazioni interattive intenzionali sempre più semplici e naturali. In generale, si assiste dunque a un crescente processo di adattamento dell'elaboratore elettronico alla natura umana. L'attenzione di questa tesi, comunque, è diretta esclusivamente alle interazioni dirette con l'elaboratore in cui il comportamento dell'utente è finalizzato a uno scopo intenzionale.

² La nuova tendenza è nota come Ubiquitous Computing (Weiser, per informazioni e bibliografia si veda <http://www.ubiq.com/hypertext/weiser/UbiHome.html>) o Reactive Environment (Buxton, <http://www.dbp.utoronto.ca/people/BillBuxton/Index.html>).

1.1. Soluzioni di avvicinamento

I computer ci aiutano a svolgere le nostre attività, ma per potere essere utilizzati necessitano impegno, conoscenze e abilità. Alla crescita della complessità di un sistema è associato l'aumento dell'impegno richiesto all'operatore. Lo sviluppo informatico, inteso come produzione di sistemi simbolici sempre più sofisticati e complessi, porta dunque con sé il problema della loro interazione con l'essere umano.

L'interazione rappresenta una situazione problematica perché i due agenti sono strutturalmente e concettualmente diversi (Norman, 1986; Hutchins, Hollan e Norman, 1986). L'utente interagisce con il computer al fine di raggiungere determinati *obiettivi* che si rappresenta in termini psicologici. Il computer, invece, comunica il suo *stato* corrente in termini fisici. Tale differenza rappresenta il principale divario (*gulf*) che deve essere superato durante l'interazione³. Teoricamente esistono due soluzioni di avvicinamento: (a) si può *addestrare* l'utente per renderlo capace di usare il sistema; (b) si può *agire* sul sistema costruendolo in modo compatibile alle caratteristiche dell'utente.

La prima alternativa rappresenta l'unica soluzione esplorata fino alla fine degli anni '70⁴. Essa era resa necessaria dalla tradizionale metodologia informatica in cui il progetto del sistema partiva dalle funzionalità da implementare senza considerare il fattore umano (approccio *centrato sul sistema, system-centered*).

L'*addestramento* tipicamente richiede lunghi e faticosi periodi di studio; inoltre non garantisce il successo. Nel caso in cui un sistema risulti troppo complesso l'utente può reagire con due comportamenti: (a) *selezione* più o meno consapevole di un sotto insieme di funzionalità su cui concentrare i propri sforzi; (b) *rifiuto* dello strumento. La *selezione* rappresenta un problema, specialmente se è involontaria. Ciò avviene spesso, non solo nell'interazione con il computer, ma anche nel caso di strumenti elettronici stereotipicamente ritenuti più semplici e familiari. Fenomeni di sotto utilizzo inconsapevole sono stati evidenziati anche dopo nove mesi di uso continuativo di un'autoradio. Eppure le *funzionalità sconosciute* erano utili, i controlli visibili e le procedure di utilizzo dettagliatamente descritte nel manuale di istruzioni (van Nes e van Itegem, 1990). Riflettendo sulla nostra esperienza personale possiamo verificare l'elevata frequenza di simili comportamenti. Si consideri, per esempio, il videoregistratore di casa e si paragoni l'elenco di tutte le funzioni conosciute con la lista delle funzioni descritte nelle istruzioni. Molto probabilmente i due insiemi non coincidono: il videoregistratore ha più funzionalità di quante ne conosciamo. Il sotto utilizzo è sintomo di errori nella

³ Dal punto di vista dell'utente, il divario si manifesta sia a livello di esecuzione di una data azione (*gulf of execution*), che a livello di valutazione del risultato (*gulf of evaluation*).

⁴ Il primo esempio di progettazione centrata sull'utente risale alla realizzazione di OMS Olympic Message System nel 1984 (Gould, Boies, Levy, Richards, e Schoonard, 1987; Preece, Rogers, Sharp, Benyon, Holland e Carey, 1994).

progettazione. Se le funzionalità sconosciute sono inutili, allora non avrebbero dovuto essere implementate in quanto potenziali distrattori. Se, invece, sono utili, allora il sistema risulta inadeguato al compito che deve supportare.

Il comportamento di *rifiuto* può limitarsi allo specifico sistema o estendersi all'intera area informatica. Il *rifiuto puntuale* rappresenta essenzialmente un problema economico. Lo specifico sistema viene respinto dal mercato. Il *rifiuto generale* ha invece conseguenze più serie. L'esclusione di una parte della popolazione dall'utilizzo di strumenti informatici, vale a dire dalla possibilità di accesso e gestione di una quota rilevante dell'informazione, può a lungo andare portare alla formazione di una nuova classe sociale a rischio: gli *emarginati elettronici* (De Angeli, 1995).

La presa di coscienza delle difficoltà legate all'addestramento ha portato all'affermazione di una nuova metodologia di progettazione: l'approccio *centrato sull'utente* (*user-centered*). Tale approccio si propone di superare il divario fra esseri umani e elaboratori elettronici realizzando *sistemi a misura d'utente* (De Angeli, 1993). A questo fine, la progettazione parte dallo studio degli effettivi bisogni dell'essere umano e continua a riferirsi ad essi attraverso tutto il processo di realizzazione (Gould e Lewis, 1985; Norman e Draper, 1986; Booth, 1989). Ciò implica il superamento del tradizionale errore nella rappresentazione dell'utente, che consisteva nell'attribuirgli un illimitato o generalizzato grado di flessibilità e adattabilità (De Angeli, 1993). Al contrario, si fa strada l'idea che un'interazione efficace debba sfruttare anche le proprietà dei comportamenti *skill-based* (Rasmussen, 1983; Reason, 1990), che per loro natura sono difficilmente modificabili. La nuova filosofia comporta un netto mutamento nella natura della progettazione che da attività esclusivamente tecnica e prevalentemente individuale tende a trasformarsi in attività interdisciplinare e d'équipe.

1.2. Verso il dialogo flessibile

Gli attuali computer sono radicalmente diversi dai primi elaboratori che quarant'anni fa cominciarono ad apparire sul mercato. La differenza non risiede solo nella configurazione fisica (si paragoni un attuale portatile con i giganti di una volta), né nelle prestazioni. Dal punto di vista dell'utente, la differenza risiede soprattutto nella possibilità di accesso alla tecnologia, vale a dire nella maggiore facilità di utilizzo. I primi calcolatori erano estremamente difficili da usare e questo per vari motivi (Preece, Sharp, Benyon, Holland e Carey, 1994): (a) il costo della macchina era talmente elevato che al suo confronto quello del fattore umano era irrisorio; (b) l'utenza era esclusivamente composta da esperti, prevalentemente ingegneri, se non addirittura dagli stessi sviluppatori; (c) non esistevano conoscenze specifiche su come semplificare l'interazione.

Comunque, fin dalle sue origini l'avvento del computer ha profondamente cambiato il quadro dei sistemi utente-macchina. Da prevalentemente *motoria* l'interazione è diventata quasi completamente *cognitiva*, coinvolgendo accanto ad aspetti mentali *periferici*, quali la percezione di stimoli, processi mentali *superiori*, quali abilità di *problem-solving* o di apprendimento (Norman e Draper, 1986; Stock, 1987a). Il progressivo adeguarsi del computer alle caratteristiche umane tende ad allargare la sfera dell'interazione includendo anche alcuni fenomeni sociali, caratteristici della comunicazione fra esseri umani.

Quando un utente interagisce con un sistema informatico ha luogo uno scambio di messaggi nei due sensi: dall'utente al computer e dal computer all'utente (Airenti, Bara e Colombetti, 1987). Per comodità esplicitativa definiamo tale scambio con il termine *dialogo*, pur essendo consapevoli che nessuno dei sistemi attuali permette lo sviluppo di un vero e proprio dialogo (§ 2.1.1). Al crescere della capacità interattiva dei sistemi utente-computer, il processo dialogico tende a trasformarsi lungo il continuum *vincolato-flessibile*.

Il dialogo altamente *vincolato* è la tipica forma di interazione fra utenti e macchine. Lo scambio è rigidamente determinato dalle capacità interattive del partner più debole: si svolge sulla base di un numero limitato di azioni che richiedono la conoscenza di una specifica sintassi e sono caratterizzate da procedure ripetitive. I gradi di libertà dell'utente sono ridotti a zero: o conosce esattamente sintassi e procedura oppure l'interazione è destinata a fallire. Prototipo di dialogo *flessibile* è invece la comunicazione fra umani. Lo scambio è caratterizzato dalla libertà e dalla creatività del linguaggio naturale; di conseguenza il suo andamento è funzione di un insieme assai vasto di variabili.

L'interazione con un sistema a dialogo vincolato rappresenta dunque un tipico atto strumentale. Il suo andamento è principalmente funzione della compatibilità fra strategie di soluzione della macchina, capacità cognitive dell'utente e difficoltà del compito. Nel caso di sistemi flessibili il comportamento umano, invece, è funzione di molte più variabili, legate alla specifica situazione interattiva e alle caratteristiche individuali dell'utente.

1.2.1. Lo sviluppo delle interfacce

Per comprendere come negli anni si sia andato evolvendo il dialogo utente-computer proviamo a ripercorrere le tappe fondamentali della storia delle interfacce, quei sistemi informatici che, codificando e traducendo l'informazione scambiata, mettono in relazione l'utente e la parte esecutiva del sistema informatico (Chignell e Hancock, 1988; Grundin, 1993). In questo seguiremo la proposta di Nielsen (1993), che identifica cinque generazioni⁵ caratterizzate da un dialogo sempre più flessibile e di conseguenza da un'utenza sempre più vasta ed eterogenea.

- *Prima generazione*, (1945-1955). L'utilizzo dei primi computer, i sistemi *batch*, non può essere definito propriamente *interazione*, ma piuttosto *operazione* su macchina (Stock, 1987b). Lo scambio veniva costretto in un unico momento temporale in cui l'utente forniva, in un codice estremamente criptico⁶, tutte le istruzioni e i dati necessari al raggiungimento del suo obiettivo. L'output era prodotto solo alla fine del ciclo di elaborazione dell'input. Il controllo e i gradi di libertà dell'utente erano quindi praticamente nulli. Per la loro estrema difficoltà i sistemi batch furono utilizzati da un'utenza estremamente limitata, composta prevalentemente da ingegneri.
- *Seconda generazione*, (1955-1965). Con i sistemi *line-oriented* si comincia a parlare di interfacce vere e proprie per quanto ancora estremamente limitate. Gli input dell'utente venivano prodotti in concomitanza a specifiche linee di comandi⁷. L'interazione era basata su un modello domanda-risposta altamente vincolato e completamente diretto dal computer. Questa procedura rende il recupero dall'errore estremamente difficile perché preclude all'utente la possibilità di modificare l'andamento del dialogo. L'utenza era composta da tecnici, informatici e ingegneri. In questo periodo comunque cominciò a delinearsi un primo interesse verso bisogni e capacità dell'utente, testimoniato per esempio dai lavori di Licklider (1960).
- *Terza generazione*, (1965-1980). Le interfacce *full-screen* permettono un'interazione che riproduce un compito di compilazione di moduli. Vari campi sono compresenti sullo schermo; Ognuno richiede parte dell'informazione necessaria all'esecuzione del

⁵ I riferimenti temporali dello schema proposto si riferiscono al momento in cui i pionieri dell'informatica hanno cominciato ad utilizzare una data generazione di interfacce. Risultano quindi inadeguati se applicati alla diffusione delle stesse sul mercato.

⁶ I programmi venivano scritti in linguaggio macchina, sequenze di 0 e 1, con perforatori di schede. Il mazzo di schede veniva poi messo su un lettore e l'output era stampato su carta.

⁷ I comandi vennero scritti inizialmente in *assembler*, un linguaggio di basso livello che manipola i registri di memoria. Più avanti furono invece utilizzati anche linguaggi di programmazione di alto livello.

comando. La scelta della sequenza di compilazione è delegata all'utente. Caratteristiche distintive di tali sistemi sono menus statici gerarchici⁸ (Paap e Roske-Hofstrand, 1988) e tasti funzione⁹ (*key-function*). L'utilità dei menus deriva da una caratteristica della memoria umana: il riconoscimento è più semplice del recupero. Il limite principale è che il passaggio da un menu all'altro nasconde l'interazione pregressa e rischia dunque di disorientare l'utente. I tasti funzione hanno il vantaggio di sveltire l'interazione, ma inevitabilmente richiedono apprendimento mnemonico. Le interfacce *full-screen* aumentarono i gradi di libertà dell'utente, che poteva scegliere fra due possibilità di scambio e aveva un minimo di influenza sull'andamento dell'interazione. Questi sistemi si diffusero particolarmente nel settore dell'informatica di gestione, dove sono tuttora utilizzati. Tipici utenti sono impiegati e manager a cui viene richiesto uno specifico periodo di addestramento, ma non approfondite conoscenze informatiche.

- *Quarta generazione*, (1980-1995). L'attuale generazione di interfacce è rappresentata dalle *interfacce grafiche* (*Graphical User Interface* o GUI), quelle su cui probabilmente molti di noi sono soliti lavorare¹⁰. Le componenti base sono finestre, icone, menu e dispositivi di puntamento¹¹ (il più tipico: il mouse). Il principale stile di interazione è la manipolazione diretta di oggetti grafici tramite mouse (Shneiderman, 1983; Hutchins, Hollan e Norman, 1985; Hutchins, Hollan e Norman, 1986). La qualità dell'interazione dipende dunque dall'*affordance* degli oggetti. È questo un concetto mutuato dalla teoria ecologica della percezione con il quale si definiscono proprietà degli oggetti ambientali relative ad un organismo (Gibson, 1979, 1982). Tali proprietà si riferiscono alle azioni che l'individuo può compiere. La conseguenza dell'applicazione del concetto di *affordance* alla manipolazione diretta del sistema è la filosofia *what you see is what you get*¹². In seguito a ogni azione la rappresentazione dell'interfaccia viene aggiornata permettendo così all'utente un continuo ed elevato controllo sull'interazione. Anche la flessibilità del dialogo risulta notevolmente

⁸ Rappresentazioni di comandi in cui ogni singolo comando è indicato da un'etichetta linguistica, selezionando la quale è possibile agire direttamente sullo stato del computer o visualizzare altri comandi.

⁹ I tasti funzione permettono di eseguire una sequenza ben definita di comandi premendo un tasto specifico cui sono associati (per esempio F1).

¹⁰ Per quanto la storia di tali sistemi possa essere fatta risalire alla prima metà degli anni '60 con i lavori di Sutherland e gli studi sul mouse di Engelbart; la loro implementazione risale solo al 1980.

¹¹ Le interfacce grafiche sono conosciute anche con l'acronimo WIMP (*Windows, Icon, Menu, Pointing device*). L'utilizzo delle finestre permette il movimento tra varie attività; le icone sono rappresentazioni grafiche di oggetti, concetti o comandi; i menu sono dinamici e possono essere di due tipi: *pop-up* (appaiono dal nulla in seguito a una specifico input del mouse) e *pull-down* (si aprono come saracinesche in seguito a una specifico input del mouse su un'apposita etichetta linguistica).

accreciuta: una stessa azione può essere eseguita in più modi a seconda delle preferenze dell'utente. Per esempio, in un sistema di videoscrittura un file può essere salvato: (a) cliccando sull'icona che rappresenta il dischetto; (b) selezionando l'item salva dal menu linguistico; (c) tramite scorciatoia (*shortcuts*) pigiando contemporaneamente due tasti specifici. La manipolazione diretta apre il mondo dei computer a una vasta utenza che, in seguito a un addestramento piuttosto ridotto, può utilizzare il computer per scopi estremamente eterogenei. Parallelamente all'avvento delle interfacce grafiche si afferma la HCI.

- *Ultima generazione*, (1995- ??). L'ultima generazione di interfacce sta cominciando in questi anni nei laboratori di tutto il mondo. Sarà la prima generazione sviluppata prevalentemente tramite l'approccio centrato sull'utente e dunque dovrebbe essere in assoluto la più compatibile con le caratteristiche umane. I cambiamenti in programma sono molteplici e agiscono su aspetti diversi dell'interazione. È possibile comunque evidenziare un comune denominatore: le dimensioni dell'interazione tendono ad aumentare e di conseguenza il dialogo diventa sempre più flessibile. Le interfacce a manipolazione diretta includono anche il movimento, nella forma di vere e proprie animazioni (Baecker, Small e Mander, 1991) o il suono (Gaver W., 1986; Blattner, Sumikawa, e Greenberg, 1989; Gaver B., 1993). La modalità di azione sugli oggetti si sta velocemente modificando verso una maggiore naturalezza. Il mouse è sostituito da penne elettroniche (Cole et al., 1995) e presto l'attivazione di icone sarà supportata presto dallo sguardo (Jacob, 1993). I sistemi di realtà virtuale e le interfacce 3D permettono di ampliare la possibilità di azione fisica. Vari sistemi di riconoscimento del linguaggio naturale sono già sul mercato e la ricerca è in un momento di particolare fervore, anche grazie alla rinuncia dell'idea di riprodurre la comunicazione fra umani (Gentner e Nielsen, 1996). In aumento sono anche le dimensioni cognitive dell'interfaccia, che viene dotata di moduli *intelligenti*, vale a dire di appropriati mezzi per comprendere le esigenze dell'utente e guidarlo nell'esecuzione del compito (Chignell e Hancock, 1988). Questo filone si sta prevalentemente orientando verso la costruzione di *agenti*, personaggi che agiscono nell'ambiente virtuale per aiutare gli utenti (Laurel, 1991). Tipicamente gli agenti sono caratterizzati da sembianze umanoidi e da un comportamento estremamente naturale. Ciò produce un aumento della dimensione emotiva dell'interazione: oltre a comunicare in linguaggio naturale, gli agenti saranno in grado di riprodurre espressioni facciali e gestuali tipiche degli stati emotivi (Cassell e Prevost, 1996). La frontiera più estrema dello sviluppo tecnologico

¹² Abbreviata WYSIWYG e pronunciato come wizzy-wig, la frase si è presto trasformata in una coppia di parole, con cui vengono identificati tutti quei programmi dove l'output del sistema rappresenta pienamente e chiaramente il risultato dell'azione dell'utente (Thimbleby, 1990).

è comunque rappresentata dai sistemi multimodali, strumenti capaci di condensare più modalità interattive in un'unica interfaccia. La sfida dei nuovi sistemi è dunque la conquista del mercato degli emarginati tecnologici.

1.3. I sistemi multimodali

I sistemi multimodali sono strumenti informatici che non soltanto supportano l'interazione attraverso più canali di comunicazione, proprietà comune a tutti i sistemi multimediali, ma che sono anche in grado di estrarre l'informazione essenziale da ciascuno di essi per usarla nel comporre un significato unico e coerente. Durante l'interazione l'utente può parlare o scrivere nella propria lingua madre, gesticolare, disegnare o attivare icone con lo sguardo. Il sistema è in grado di *comprendere* il messaggio indipendentemente dalla modalità in cui viene trasmesso e di fornire vari tipi di output, parlato, scritto o tattile, che entro una certa misura risultano equivalenti per l'essere umano, esempio paradigmatico di sistema multimodale.

Da un punto di vista tecnologico la multimedialità è ormai raggiunta. La multimodalità invece risente delle disomogeneità tra il livello di sviluppo delle varie modalità (Arens e Hovy 1990; Mackinlay, Card e Robertson, 1990; Coutaz, Nigay e Salber, 1993; Ole Bersen, 1993). Accanto a modalità ampiamente analizzate, quali la grafica bidimensionale, esistono modalità ancora in fase di sperimentazione, quali il linguaggio naturale scritto o parlato, e altre, quali la gestualità, dove gli studi sono appena cominciati (Reithinger, 1987; Schmauks, 1987, Schmauks e Reithinger, 1988, Schmauks e Willie, 1991; De Angeli *et al.*, 1996; Oviatt, 1996; Oviatt *et al.* 1997). Inoltre, mentre ci sono stati dei progressi significativi nell'utilizzo delle singole modalità, meno enfasi si è avuta nell'integrazione e nella coordinazione di più canali.

La multimodalità, oltre ad agevolare l'utente, tende ad agevolare il computer. I primi tentativi di implementazione derivano dall'esigenza di rendere più robusti gli attuali sistemi di riconoscimento del linguaggio naturale. La disponibilità di più canali tende infatti ad aumentare il rapporto segnale-rumore, ossia la quantità di informazione utilizzabile dal sistema per capire le intenzioni comunicative dell'utente (Maybury, 1993; Cole *et al.*, 1995). L'interazione multimodale appare superiore a quella unimodale sia nella comunicazione diretta di un dato significato, *livello dialogico*, sia nell'interpretazione del dialogo, *livello meta-dialogico*. A livello dialogico la disponibilità di più canali permette di semplificare la formulazione linguistica di alcuni contenuti (De Angeli *et al.*, 1996; Oviatt, 1996; Oviatt *et al.*, 1997). È questo il caso per esempio dell'identificazione di referenti nello spazio, in cui un gesto può sostituire complicate espressioni verbali (§ 6 e 7). A livello meta-dialogico, invece, l'utilizzo di più modalità facilita la comunicazione delle capacità comunicative del computer e di conseguenza dirige il comportamento dell'utente

verso tali capacità. Un grafico può comunicare la coesione che il sistema informatico attribuisce alle parti del dialogo in modo più diretto e agevole di un atto linguistico. La manipolazione diretta può consentire un più agevole recupero dall'errore, evitando complessi e noiosi sotto-dialoghi di chiarimento (Zancanaro, Stock e Strapparava, 1993; Stock, Strapparava e Zancanaro, 1995; Stock, 1995; Zancanaro, Stock e Strapparava, 1994; Zancanaro, Stock e Strapparava, in stampa).

1.3.1. Interfacce multimodali

Nell'architettura di un sistema multimodale l'interfaccia assume un ruolo fondamentale. Tale modulo deve mediare l'interazione in senso ampio, coordinando e dirigendo le modalità comunicative adottate dai due partner verso livelli di comprensione ottimali (Walhlster, 1992). La qualità dell'interazione è dunque funzione del grado di flessibilità e adattabilità dell'interfaccia. Essenzialmente tale modulo deve permettere il costante processo di mutuo accomodamento tipico della comunicazione fra umani, per cui il parlante tende a modificare i suoi atti comunicativi sulla base del contenuto trasmesso, delle caratteristiche dell'ascoltatore e del contesto in cui lo scambio ha luogo (Grice, 1975).

Flessibilità e adattabilità di un'interfaccia multimodale possono essere differenziate sulla base di tre livelli crescenti di complessità. Il primo, *il livello delle risorse fisiche*, implica che il sistema sia in grado di adattare la presentazione di un contenuto in funzione della specifica modalità comunicativa utilizzata, ottimizzando così le caratteristiche intrinseche di ogni canale di comunicazione. Il secondo, *il livello delle preferenze dell'utente*, implica che il sistema adegui il suo output sulla base della modalità di input dell'utente e che renda più facilmente disponibili le modalità di interazione preferite dall'utente. L'ultimo e più complesso livello di adattabilità è *il livello dei contenuti*. L'interfaccia è considerata adattabile se è in grado di scegliere la modalità di comunicazione più adatta a trasmettere un dato contenuto.

Lo sviluppo di interfacce multimodali è un compito molto complesso, il cui successo è subordinato alla disponibilità di modelli predittivi della comunicazione che spieghino i processi cognitivi, percettivi e motori sottostanti l'uso coordinato ed efficace di varie modalità di I/O, nonché i processi sociali sottostanti la scelta di una determinata strategia interattiva. Le conoscenze relative all'interazione multimodale attualmente disponibili sono molto scarse (Oviatt, 1996; Oviatt *et al.*, 1997). Tuttavia si sta assistendo a un crescente interesse verso l'argomento, sia per obiettivi pratici sia perché lo sviluppo dei sistemi multimodali rappresenta un'attraente sfida per la HCI. Infatti, esso è reso possibile in un momento in cui la disciplina può vantare una sostanziosa esperienza e una vitalità impensabile in passato. Allo stesso tempo però l'argomento richiede una profonda

riflessione critica per adattare ai nuovi sistemi flessibili metodi e teorie sviluppati per lo studio di sistemi vincolati.

Dall'esperienza acquisita con i sistemi di riconoscimento del linguaggio naturale è maturata la consapevolezza che la comunicazione fra umani non rappresenta un modello adeguato dell'interazione utente-computer. Nel caso dei sistemi multimodali l'inadeguatezza viene esasperata dalla possibilità di unire modalità che in natura non co-occorrono (§ 6 e 7). Di conseguenza la ricerca si sta orientando prevalentemente verso un approccio empirico, basato principalmente sulla simulazione (§ 3.5).

1.4. Artefatti socio-cognitivi

Il progressivo adattamento del computer all'utente riduce il divario fra le due entità e quindi trasforma l'elaboratore in qualcosa di diverso da una macchina, in un'entità più compatibile all'essere umano. La flessibilità dell'interazione tende a estremizzare la *concezione antropomorfa* che ha caratterizzato il computer fin dai suoi albori¹³ (Brighetti, Contento e Tabossi, 1987; De Angeli, 1991). Di conseguenza da artefatto¹⁴ cognitivo il computer si trasforma in un artefatto socio-cognitivo.

Questa affermazione può sembrare curiosa, ad alcuni addirittura provocatoria. Se infatti, la definizione di computer come artefatto cognitivo (Norman, 1991) è ampiamente accettata l'estensione di tale concetto alla sfera sociale è un'idea recente e controversa. Le condizioni sottostanti al suo concepimento potevano realizzarsi solo dato l'attuale stadio di sviluppo tecnologico e la piena applicabilità della definizione sarà possibile solo in futuro, quando gli attuali prototipi saranno diventati sistemi di pubblico dominio. Proviamo dunque ad analizzarla nel dettaglio.

Norman definisce come artefatti cognitivi quegli strumenti artificiali che immagazzinano, forniscono o manipolano informazioni. In questo senso un libro, la lista della spesa, la logica o la matematica sono artefatti cognitivi, oggetti creati dall'essere umano per adempiere a funzioni rappresentazionali e migliorare¹⁵ le prestazioni umane nell'esecuzione di compiti. Il computer è una macchina che elabora simboli. La sua capacità di svolgere operazioni logiche o matematiche e di memorizzare informazioni a una velocità superiore e senza i limiti quantitativi propri dell'essere umano, lo hanno ben

¹³ Si pensi per esempio all'uso dei termini *memoria* e *linguaggio* per indicare alcune caratteristiche del computer o al più recente aggettivo *amichevole* con cui si denota la qualità dell'interfaccia.

¹⁴ In questo contesto, la parola *artefatto* è utilizzata in funzione di sostantivo con il significato di *manufatto*. In tale accezione il termine è un neologismo derivato per calco dall'inglese *artifact*. In italiano la parola esisteva già, ma in funzione di aggettivo e con il significato di *adulterato, insincero* (Devoto e Oli, 1990).

¹⁵ Nella definizione originale l'autore utilizza il verbo *affect* (influenzare). Ho preferito il termine *migliorare*, perché ritengo che un artefatto cognitivo abbia ragione di esistere solo se accresce le capacità umane, intese come potenzialità cognitive, edonistiche o ludiche, nell'esecuzione di compiti.

presto reso il prototipo della categoria artefatti cognitivi. Si ricordi che l'idea stessa di essere umano come *agente cognitivo* deriva dal paragone con il calcolatore elettronico (per esempio, Miller, Galanter e Pribram, 1960).

La resistenza nei confronti della definizione di computer come artefatto sociale risiede innanzitutto nella naturale sensazione di disagio che deriva da un ossimoro poco familiare. Il dissenso è inoltre accentuato dallo specifico contrasto logico insito nell'accostamento dei due termini, che tende a sollevare profonde problematiche filosofiche, ontologiche e addirittura religiose. Il termine artefatto identifica i prodotti derivati da lavorazione a mano o a macchina. Dunque l'artefatto è per sua natura un prodotto *artificiale*. Al contrario, l'aggettivo sociale è usato per qualificare agenti o attività *naturali*. Di conseguenza, il concetto di artefatto sociale estende una qualità esclusiva di entità naturali a un'entità artificiale. Ciò implica il progressivo indebolimento della radicale dicotomia naturale/artificiale e un passo avanti verso la definizione del continuum naturale-artificiale, che gli esseri umani, specie egemone della categoria naturale, sentono come una minaccia al loro predominio.

Le tecnologie interattive maggiormente responsabili della trasformazione del computer in artefatto socio-cognitivo sono i sistemi di riconoscimento del linguaggio naturale e in maniera ancora più decisiva i sistemi multimodali. Nel momento in cui il linguaggio entra nell'interazione vengono attivati una serie di meccanismi che guidano il suo utilizzo nella normale situazione d'uso: la conversazione fra umani (De Angeli, 1991; Naas, Steuer, e Tauber, 1994; Naas, Steuer, Herinksen e Dryer, 1994). Dalla voce e dal modo di esprimersi all'artefatto può essere attribuito un carattere e un'identità: di conseguenza l'interazione viene influenzata dagli stereotipi categoriali dell'utente. La rappresentazione mentale che un utente occasionale si forma dell'interlocutore meccanico assomiglia più alla percezione di personalità di un partner che alla percezione fisica di uno strumento artificiale. Nell'interazione tendono a essere rispettate delle norme sociali tipiche della comunicazione fra umani, quali alcuni vincoli di gentilezza e la differenziazione sé-altro.

1.5. *Difficoltà di analisi dell'interazione flessibile*

La flessibilità degli artefatti socio-cognitivi rende l'interazione più semplice per l'utente, ma allo stesso tempo più complessa ai fini dell'analisi. Un'interazione *vincolata* è facilmente rappresentabile, mentre la formalizzazione di un dialogo flessibile è molto complessa, se non impossibile. Analizziamo un esempio pratico, paragonando il processo dialogico richiesto da un'interfaccia *line-oriented* per l'esecuzione di una tipica azione informatica (copia di un file su un dischetto nuovo), con quello permesso da un'interfaccia multimodale.

Nel primo caso consideriamo l'interazione in ambiente MS-DOS. L'utente ha a disposizione un'unica procedura rigidamente determinata dal sistema e che richiede una conoscenza specifica (vedi Finestra 1). Il dialogo è estremamente limitato: il suo andamento può essere previsto con facilità sulla base delle conoscenze e delle capacità cognitive dell'utente.

Se però il sistema operativo viene mediato da un'interfaccia multimodale¹⁶, che permette l'interazione tramite linguaggio parlato e manipolazione diretta, le possibilità aumentano molto. Innanzitutto l'utente è libero di raggiungere il suo obiettivo in un unico o in più turni conversazionali. Inoltre, per ogni sotto azione sono disponibili numerose alternative a seconda delle modalità utilizzate, delle preferenze dell'utente e del contesto in cui l'interazione ha luogo. Il comando di formattazione potrebbe essere trasmesso tramite linguaggio e puntamento indicando il dischetto e dicendo *formattalo, inizializzalo, scrivigli sopra* o in tanti altri modi ancora. Si potrebbe inoltre utilizzare solo il linguaggio parlato (*inizializza il dischetto*), solo il linguaggio scritto (*form, format*¹⁷) o solo la manipolazione diretta (selezionando il dischetto e poi l'item formattazione dall'apposito menu). Aumentando il numero di possibilità corrette cresce la possibilità di successo dell'azione dell'utente e diminuisce la possibilità di prevedere la forma sintattica precisa dell'interazione.

¹⁶ È questo un esempio di pura immaginazione. Attualmente non esistono sistemi di questo tipo, né a quanto mi risulta sono in fase di sviluppo prototipi simili. L'interfaccia può essere immaginata simile a quella disponibile in ambiente WINDOWS per la gestione delle risorse del computer.

¹⁷ Un recente studio ha dimostrato che gli utenti quando usano una penna elettronica tendono a preferire l'uso di abbreviazioni (Oviatt.e.Olsen, 1994).

C:\>format a:

Inserire un nuovo disco per l'unità A:
e premere invio quando pronti...

(INVIO)

Controllo del formato del disco.

Verifica di 1,44 KB in corso

Formattazione completata.

Etichetta di volume (11 caratteri, INVIO per non darne)?

disco A

{il sistema riporta le caratteristiche del dischetto formattato, spazio totale ecc.}

1,457,664 byte di spazio totale su disco

1,457,664 byte disponibili su disco

512 byte in ogni unità di allocazione.

2,847 unità di allocazione disponibili su disco.

Numero di serie del volume: 1D62-07DB

Formattare un altro disco (S/N)?

N

C:\> copy [nome file] a:

1 file copiato

Finestra 1. Esempio di dialogo vincolato fra un utente e un'interfaccia *line-oriented*.

Procedura per la formattazione e la copia di un file su dischetto in ambiente MS-DOS. Gli scambi prodotti dall'utente sono riportati in grassetto.

2. HCI: una disciplina dinamica

La prima sensazione di un ricercatore che si avvicina alla HCI è probabilmente un misto di stupore e smarrimento. La piacevole sorpresa deriva dalla straordinaria quantità di materiale bibliografico, congressi, siti internet e mailing list attualmente disponibili su quest'area di ricerca tanto giovane. Si consideri, per esempio, che la bibliografia essenziale presentata da Nielsen (1993) nel suo saggio *Usability Engineering* occupa 20 pagine e descrive almeno una decina di riviste specialistiche, svariati manuali, una trentina fra libri e raccolte di articoli, nonché varie videocassette, conferenze e raccolte bibliografiche su supporto cartaceo o informatico. Navigando in INTERNET si può accedere a otto siti specifici che descrivono più di un centinaio di laboratori o progetti di ricerca¹⁸. Le quattro mailing list gestite dallo *XEROX Park* spediscono giornalmente avvisi di conferenze, richieste di articoli e offerte di lavoro.

Il piacere che deriva dalla scoperta di tanta vitalità scientifica è però presto affiancato dalla difficoltà di selezione e fruizione del materiale. Il campo è infatti caratterizzato da un intrinseco disordine, prezzo forse inevitabile di una tanto rapida evoluzione, e dalla naturale complessità di un settore interdisciplinare. Una delle maggiori sorgenti di confusione deriva da una generalizzata incoerenza linguistica che riflette ed esaspera il mancato accordo sulla definizione di oggetto e metodologie di studio.

Tanto disordine è principalmente dovuto alla natura *dinamica* dell'unità di analisi, che deriva dallo stretto rapporto fra HCI e sviluppo tecnologico (Carroll e Campbell, 1986; Carroll, Kellog e Rosson, 1991; Myers, 1996). La disciplina nasce dal progresso informatico al preciso scopo di fornirgli linfa vitale. In questo *loop* ogni potenziamento della componente informatica implica necessariamente un allargamento dell'oggetto di studio. Il cambiamento nella natura dell'interazione legato allo sviluppo dei sistemi interattivi flessibili è solo uno fra i numerosi esempi di tale fenomeno (§ 1.2). Si ricordi anche l'avvento della tecnologia delle reti informatiche, che ha esteso il dominio di studio dall'analisi dell'interazione fra *un* computer e *un* utente, all'analisi dell'interazione fra *più* computer e *più* utenti (De Angeli, 1995).

Per facilitare i primi passi del nostro ricercatore, in questo capitolo viene presentato un quadro generale della HCI, che sottolinea da più punti di vista il suo intrinseco

¹⁸ Un utile indirizzo di partenza è <http://www.ida.liu.se/~miker/hci/labs/>. La pagina in questione è gestita da Mikael Ericsson presso il *Department of Computer and Information Science* at Linköping University in Svezia.

carattere *dinamico*. In esso vengono proposti vari tentativi di ricostruzione storica, compito estremamente insidioso. L'andamento della HCI, infatti, è caratterizzato da una storia complessa e confusa che varia a seconda della prospettiva da cui viene descritta e dei diversi ambienti in cui la disciplina si è sviluppata. Una rappresentazione temporale è dunque possibile solo nella piena consapevolezza di operare una fondamentale trasformazione sul fenomeno stesso, sovrapponendogli una linearità che il passato non ha avuto. Nonostante questa inevitabile riduzione, rimango convinta che questo percorso possa aiutare a comprendere che cosa è e quali siano le prospettive di ricerca della HCI.

2.1. *Human-Computer Interaction*

Una delle maggiori difficoltà interne alla HCI deriva da una diffusa incoerenza linguistica. Il problema è inoltre complicato dal frequente ricorso a neologismi (un esempio fra tanti: il termine italiano *artefatto*), dall'abitudine tipicamente informatica dell'uso di acronimi (solo per ricordare quelli già incontrati nell'introduzione GUI, WIMP o WYSIWYG), nonché dall'abituale utilizzo di termini specifici mutuati dalle singole discipline che contribuiscono allo studio.

Pur ammettendo che l'*instabilità terminologica* possa essere considerata una normale caratteristica dei momenti di fondazione di nuove discipline (Bara, 1984), nel contesto della HCI alcuni fattori specifici hanno contribuito a renderla più forte e problematica. I principali sono: (a) la natura dinamica dell'oggetto di studio che richiede un costante aggiornamento del vocabolario; (b) l'eterogeneità delle competenze coinvolte (§ 2.3); (c) la varietà delle origini e degli ambienti in cui la disciplina si è andata affermando. Per comprendere quest'ultimo punto si ricordi che la HCI si è sviluppata contemporaneamente dalla tradizione ergonomica e da esigenze specifiche dell'informatica, disciplina originariamente estranea all'ergonomia. Inoltre, le sostanziose ricadute applicative delle ricerche (Myers, 1996) hanno spinto alla costituzione, accanto alla comunità scientifica pubblica, di numerosi centri di ricerca industriali, la cui esigenza di riservatezza tende a ostacolare la diffusione delle conoscenze e di conseguenza la stabilità del linguaggio.

Negli anni fenomeni di instabilità terminologica hanno investito addirittura il nome della disciplina che, a seconda delle preferenze dei singoli gruppi di ricerca, venne ed è tuttora identificato da un numero sconcertante di alternative (Card, Moran e Newell, 1983; Booth, 1989; Nielsen, 1993). Tra queste, alcune nacquero come sinonimi, mentre altre vennero proposte allo scopo di sottolineare precise scelte teoriche e metodologiche. Con il tempo, però, la carenza di definizioni e la generale mobilità linguistica hanno portato alla sottovalutazione delle distinzioni semantiche, lasciando il campo nel più assoluto disordine.

Parte dei primi studi relativi al comportamento di interazione fra utenti e computer vennero prodotti dalle due discipline che tradizionalmente si occupavano dello studio dell'interazione utente-macchina: i *Fattori Umani* e l'*Ergonomia*¹⁹. Tuttora, le maggiori conferenze, riviste o libri di tali settori ospitano sessioni speciali dedicate alla HCI. Il rapido sviluppo della tecnologia informatica e la crescente attenzione dedicata all'argomento hanno portato all'affermazione di settori specifici denominati rispettivamente: *Human Factors in Computing Systems* e *Cognitive Ergonomics*. La differenza fondamentale fra i due settori e le discipline madri consiste in un totale superamento dell'originario interesse fisiologico e nella parallela affermazione di un approccio cognitivista. Infatti, per il suo utilizzo il computer pone relativamente pochi vincoli al *corpo* dell'utente, ma molti alla sua *mente*.

Già a partire dagli anni '70, prima dunque dell'affermarsi dell'interesse verso il sistema utente-computer, una diffusa insoddisfazione per il ruolo subordinato cui l'ergonomia tendeva a essere relegata dall'ingegneria (si veda per esempio, Newell e Card, 1985) indusse vari ricercatori a identificare il loro lavoro in modo autonomo. Furono proposti allora termini quali *Man-Machine Interaction* (MMI) o, per evitare valenze sessiste, *Human-Machine Interaction* (HMI) che presto vennero utilizzati anche per designare studi relativi a sistemi informatici. In teoria l'utilizzo del termine *Machine* dovrebbe contraddistinguere un campo di ricerca più vasto rispetto allo specifico dominio della HCI. In pratica, però, tale distinzione non viene rigidamente applicata (Booth, 1989), sia per la disomogeneità linguistica e storica del settore, sia per una sempre maggiore intercambiabilità tra i due campi di ricerca. Se è banale evidenziare che il computer altro non è che un esemplare della categoria macchina, è invece più interessante notare come il progresso tecnologico stia contribuendo all'attenuazione delle loro differenze. Sempre più spesso infatti le *macchine* includono sistemi di elaborazione dell'informazione nonché interfacce informatiche (§ 1).

Un termine alternativo è *Cognitive Engineering*, proposto da Norman, (Norman e Draper, 1986) per indicare un'applicazione delle scienze cognitive, il cui scopo è l'utilizzo di conoscenze scientifiche nella progettazione e costruzione di macchine. Secondo l'autore, la disciplina rappresenta un campo di ricerca autonomo che coinvolge ma non si

¹⁹ La differenza fra i due termini riflette le loro origini geografiche. Il primo nacque e si diffuse, parallelamente ai sinonimi *Human Factors Engineering* e *Human Engineering* negli Stati Uniti d'America durante la seconda guerra mondiale, mentre nel vecchio continente veniva preferita la parola *Ergonomics* (Re, 1995). In Italia, quest'ultima fu inizialmente tradotta *Ergotecnica* e poi *Ergonomia* (Bagnara e Visciola, 1987). Per quanto sia possibile identificare alcune lievi differenze fra le due discipline, in questo lavoro i termini verranno considerati sinonimi. L'ergonomia si propone come scienza globale dell'attività lavorativa: il suo oggetto di studio è la progettazione per uso umano di apparecchiature, ambienti di lavoro e procedure per l'esecuzione di compiti (Re, 1995; McCormick e Sanders, 1992). L'obiettivo è adattare le condizioni di lavoro alle capacità psico-fisiologiche del lavoratore per aumentare la produttività e garantire il rispetto dei valori umani negli ambienti di lavoro. Lo studio dell'interazione utente-macchina rappresenta dunque solo l'aspetto legato alla dimensione tecnologica del lavoro.

identifica completamente né con la psicologia cognitivista, né con le scienze cognitive e neppure con l'ergonomia (per un'introduzione al settore si vedano anche Rasmussen, 1986 o Woods e Roth, 1988). L'ingegneria cognitiva studia il comportamento umano in ambienti complessi sulla base di un approccio ecologico. Oggetto di studio sono le strategie di *problem-solving* di compiti complessi adottate da utenti e macchine. Lo studio dell'interazione utente-computer è dunque uno degli argomenti paradigmatici della disciplina.

A partire dagli anni '80, in corrispondenza all'effettivo diffondersi dell'interesse verso i sistemi informatici, furono proposti numerosi termini specifici quali HCI, oppure *Computer-Human Interaction* (CHI), *Human-Computer Communication* (HCC), *Man-Machine Communication* (MMC) o *Man-Machine Dialogue* (MMD). La posizione dei termini *Human* e *Computer* riflette l'importanza relativa attribuita ai due partner²⁰. La preferenza per i termini *communication* o *dialogue* rispetto ad *interaction* riflette invece una precisa scelta teorica tipica dell'intelligenza artificiale, che riconosce al computer maggiori capacità comunicative e una forma di intelligenza.

2.1.1. I motivi della scelta

La scelta dell'espressione *Human-Computer Interaction* in questa tesi riflette la tendenza prevalente del settore nonché precise scelte teoriche alla base dell'*Approccio Globale* (§ 4). La HCI è considerata un particolare campo dell'ergonomia cognitiva, ristabilendo così le origini naturali, spesso sottovalutate se non direttamente rifiutate. L'idea di considerarla come uno specifico settore si basa sul presupposto che il computer sia una macchina di tipo speciale, per il cui studio metodi e teorie ergonomiche sono necessari, ma non sufficienti.

La peculiarità del computer è legata a reali caratteristiche dello strumento e al modo in cui viene percepito dagli utenti. In generale il termine macchina può essere appropriato a definire lo hardware, ma sicuramente non il software. Da un punto di vista oggettivo l'elaboratore elettronico è diverso dalle macchine tradizionali (non informatiche) poiché non opera direttamente sullo stato esterno del mondo, monitorandolo o producendo energia o materia²¹. Il computer elabora informazione e quindi opera sullo stato mentale dell'utente. Al di là di queste differenze, la peculiarità del computer dipende in maniera

²⁰ Attualmente, *Human-Computer Interaction* è l'alternativa più frequente. Tuttavia uno dei più importanti eventi scientifici del settore è la conferenza CHI e la rivista edita dalla ACM (*Association for Computing Machinery*) titola *acm transactions on Computer-Human Interaction*. ACM è la principale società informatica Americana, da sempre impegnata nel campo delle interfacce di cui gestisce una delle più importanti associazioni: la ACM SIGCHI (*Special Interest Group on Computer Human Interaction*).

²¹ Questa differenza è comunque destinata a scomparire a seguito dei lavori prodotti nel settore *Ubiquitous Computing* (§ nota 2, pag. 8).

cruciale dalla nostra esperienza con esso. I dati esperenziali più significativi sono la varietà di funzioni cui può essere adibito e la sua opacità di funzionamento (Brighetti, Contento, Tabossi, 1987). Consideriamo questo aspetto tramite un esempio pratico. Se un amico ci presta l'automobile, dopo pochi minuti necessari alla localizzazione fisica della strumentazione di bordo, siamo in grado di guidarla. Anche se non siamo dei meccanici, solitamente siamo capaci di prevedere l'effetto di ogni nostra azione sulla guida. Se usiamo un nuovo sistema informatico invece abbiamo molte più difficoltà. Anche se siamo degli informatici, non possiamo prevedere con esattezza l'effetto delle nostre azioni su un sistema sconosciuto. La difficoltà di inferenza causale diminuisce la percezione di controllo, inducendo così ad attribuire al computer capacità autonome. Quanto più le capacità interattive del computer si sviluppano tanto più questa rappresentazione diventa stabile.

Il computer viene differenziato dalle altre macchine in quanto più simile all'essere umano (Brighetti *et al.*, 1987; De Angeli, 1991), eppure differenze estremamente marcate lo separano da tale polo di riferimento. Vari studi hanno dimostrato la peculiarità del dialogo utente-calcolatore anche in sistemi flessibili capaci di utilizzare il linguaggio naturale isolatamente o in concomitanza ad altri canali interattivi (Hauptmann e Rudnick, 1988; Jöhnsson e Dählback, 1988; Kennedy, Wilkes, Elder e Murray, 1988; De Angeli, 1991, De Angeli *et al.*, 1996). Le differenze principali riguardano: (a) la grammatica, che deve essere rigorosamente rispettata nell'interazione con il computer, mentre è solitamente sottovalutata nella comunicazione fra umani; (b) la ricchezza comunicativa, che è sensibilmente inferiore nell'interazione con il computer, dove le frasi sono brevi e semplici, la scelta lessicale e l'uso di riferimenti anaforici ridotti al minimo, buona parte dei canali non verbali ignorati; (c) il recupero dagli errori e la negoziazione di significati comuni, estremamente complessi nell'interazione con il computer, spontanei nella comunicazione fra umani.

La differenza fra *interazione* utente-computer e *comunicazione* fra umani è prevedibile anche su base teorica. Il concetto di comunicazione in psicologia si configura infatti come *un fenomeno articolato e complesso, un gioco sociale e primario, intrinsecamente connesso alla natura umana* (Anolli e Ciceri, 1995, pag.21). Postulato fondamentale della definizione è la reciproca attribuzione di intenzione comunicativa fra i due partner. Per quanto in ogni artefatto si possano riconoscere le intenzioni comunicative del progettista, difficilmente si giunge ad attribuire al computer autonoma e piena volontà comunicativa. L'interazione si svolge in funzione di specifiche intenzioni dell'utente e solo in domini nettamente delimitati (Airenti, Bara e Colombetti, 1987).

Inoltre, la conversazione fra umani si basa sullo scambio di segni provenienti da più canali, verbali e non. L'attribuzione di significato è subordinata alla determinazione di adeguate relazioni strutturali fra segni trasmessi da canali diversi, mediante interpretazione

sintattica, semantica e pragmatica. In futuro probabilmente i computer saranno in grado di riconoscere la funzione sintattica e comprendere il significato semantico di un insieme molto vasto di segni. L'evoluzione dell'abilità pragmatica, capacità fondamentale per l'effettiva comprensione del significato, appare invece più problematica, poiché richiede la piena comprensione del contesto in cui il significante viene prodotto e della persona che lo produce. Anche gli artefatti socio-cognitivi sono destinati a rimanere a lungo dei partner di natura speciale, umanoidi forse, ma non umani.

2.2. *Oggetto di studio*

Trovare in letteratura una definizione unanimamente accettata e soddisfacente dell'oggetto di studio della HCI può essere difficile (Preece *et al.*, 1994). Nuovamente all'origine di questo fenomeno si collocano le caratteristiche peculiari della disciplina, prima fra tutte la sua natura dinamica. Ulteriore causa è il preponderante orientamento applicativo di studi e ricerche (Myers, 1996) che ha scoraggiato riflessioni ontologiche a favore di studi pratici. La maggior parte della vasta bibliografia disponibile si occupa di aspetti funzionali dell'interazione, ma non si preoccupa di definirla globalmente, facilitata dalla valenza auto esplicativa del nome della disciplina. La carenza di definizioni è anche dovuta all'intrinseca difficoltà che risiede nella descrizione delle caratteristiche specifiche di un processo complesso (§ 4).

Quando fornite, le definizioni tendono a essere parziali; si soffermano su alcuni aspetti dell'unità di analisi e di conseguenza ne sottovalutano la complessa natura interdisciplinare. Questi tentativi soffrono di una visione parziale dell'interazione utente-computer assunta a definizione assoluta. Secondo Norman (1991), infatti, l'interazione può essere descritta in modo diverso a seconda della prospettiva da cui il sistema *persona-artefatto* viene osservato. Un osservatore esterno tende ad adottare la *prospettiva sistemica*. Da questo punto di vista egli percepisce la struttura totale *persona + artefatto* che esegue un dato compito. Tale esperienza spinge a inferire che il computer agisca direttamente sull'utente amplificandone le capacità cognitive. Un osservatore interno, invece, guarda la situazione dalla parte dell'attore, secondo la *prospettiva personale*. La sua attenzione è prevalentemente diretta verso il compito. Così, il computer viene visto come un modo alternativo di eseguire compiti mnemonici o di risolvere problemi.

Consideriamo adesso alcune delle definizioni fornite in letteratura. Questa rassegna non ha alcuna pretesa di completezza, ma serve piuttosto a dare un'idea generale della natura dinamica della HCI e a definire i precedenti teorici da cui si è sviluppato l'*approccio globale*.

2.2.1. Alcune proposte

La prima definizione rilevante ai fini dell'approccio globale è quella sviluppata, secondo la *prospettiva sistemica*, da Baecker e Buxton (1987, pag. 40). Tali autori descrivono la HCI come l'insieme di *processi, dialoghi e azioni attraverso i quali l'utente interagisce con il computer*. L'obiettivo viene identificato nell'applicazione sistematica di conoscenze relative a scopi, capacità e limiti umani, nonché a capacità e limiti del computer, all'implementazione di nuovi sistemi che permettano di diminuire il divario fra le due entità. Una decina di anni fa, dunque, il comportamento di interazione veniva definito e studiato in puri termini di scambi informativi prodotti dai due agenti.

Una fra le proposte più dettagliate e complete è quella di Booth (1989) che definisce l'oggetto di studio descrivendo cinque aree di ricerca sottostanti la HCI. Secondo la sua proposta tali settori sono i seguenti.

- *Studi relativi a hardware e software interattivi*. I lavori in questo settore sono finalizzati allo sviluppo di strumenti tecnologici che semplifichino e rendano più intuitivo l'uso del calcolatore. Tale obiettivo è subordinato alla comprensione del modo in cui diverse modalità di interazione agiscono sul comportamento dei due agenti. Gli studi si rivolgono a sistemi, prototipi e anche a nuove prospettive che lo sviluppo tecnologico permette di intravedere.
- *Studi di compatibilità fra modelli*. Le ricerche in questo settore riguardano una problematica fondamentale sottostante l'interazione, vale a dire la necessità di implementare modelli del compito adeguati a quelli posseduti dagli utenti. Esse sono dunque rivolte allo sviluppo di teorie e strumenti per la formalizzazione di modelli mentali generici istanziabili su di una classe di compiti, come per esempio modelli per la video scrittura. Queste conoscenze devono essere comunicate ai tecnici della progettazione per una loro applicazione pratica.
- *Studi a livello del compito*. Quest'area tende a sovrapporsi a quella precedente in quanto in entrambe l'attenzione è focalizzata sul compito. La differenza cruciale è però che la precedente si occupa di studiare come l'utente esegue un dato compito, mentre questa si concentra sull'analisi della natura globale del compito e delle effettive necessità che l'essere umano si trova ad avere per eseguirlo. L'obiettivo è lo sviluppo di metodologie per la comprensione dei bisogni degli utenti.
- *Studi relativi alla progettazione*. Le ricerche in questo settore sono rivolte allo sviluppo di teorie e strumenti di progettazione che permettano e stimolino la realizzazione di sistemi sempre più efficienti e adeguati all'utenza cui sono destinati. Lo scopo finale è il passaggio dalla progettazione centrata sul sistema a quella centrata sull'utente.

- *Studi relativi all'impatto sulle organizzazioni.* Le ricerche in questo settore sono rivolte allo studio degli effetti che l'introduzione di nuove tecnologie ha su individui, gruppi o società. Lo scopo è la prevenzione dei conflitti sociali e la salvaguardia dei fondamentali valori umani nel lavoro informatico.

La definizione di Booth è sicuramente di grande interesse. Il principale elemento innovativo è l'assunzione della *prospettiva personale* che determina l'introduzione del *compito*, come diretto oggetto di analisi. Tuttavia, per quanto l'autore si proponga di ottenere esattamente il risultato opposto, la sua idea rischia di aumentare la divisione settoriale che caratterizza e, secondo l'approccio globale, ostacola la comprensione dell'interazione utente-computer (§ 4).

Più recentemente, Preece e colleghi (1994, glossario pag. 714) estendono l'originaria proposta di Baecker e Buxton (1987) in questo modo: la HCI è lo *studio di processi, dialoghi e azioni che l'utente adotta nell'interazione con un computer in un dato ambiente*. La novità sostanziale è l'inclusione dell'ambiente (in cui viene implicitamente inserito anche il compito) come diretto elemento di studio (si veda anche ACM SIGCHI, 1992). Booth, infatti, considerava il contesto come un'entità prevalentemente passiva che può venire modificata dall'introduzione di nuove tecnologie e non come un'entità attiva che può a sua volta influenzarne l'utilizzo. Nella nuova visione, gli scopi della HCI vengono identificati nel miglioramento di sistemi interattivi che comprendono il computer tramite: (a) comprensione dei fattori psicologici, ergonomici, organizzazionali e sociali sottostanti il comportamento dell'utente; (b) sviluppo di strumenti e tecniche che permettano di applicare tali conoscenze alla progettazione; (c) realizzazione di nuovi sistemi più adeguati alle caratteristiche dell'utente.

Riassumendo, negli anni l'oggetto di studio è andato via via allargandosi, includendo prima il compito e poi l'ambiente. Costante nel tempo è rimasta invece l'idea di una disciplina con forti ricadute applicative, il cui scopo ultimo è la realizzazione di nuovi sistemi interattivi. Inoltre, tutte le definizioni accettano che l'oggetto di studio per sua stessa natura cada al di fuori delle tradizionali ripartizioni fra aree di ricerca e richieda quindi un approccio di tipo interdisciplinare.

2.3. *L'approccio interdisciplinare*

Per quanto la necessità di un approccio interdisciplinare sia un'opinione ampiamente condivisa, il numero delle discipline ritenute salienti varia a seconda della personale preferenza degli autori e tende a crescere parallelamente allo sviluppo della disciplina stessa. Per esempio, Newell e Card, (1985) evidenziano quattro aree di studio: computergrafica, fattori umani, psicologia cognitiva e intelligenza artificiale. Booth (1989) ne elenca dieci: ingegneria del software, linguistica computazionale, intelligenza artificiale, scienze cognitive, sociologia, ergonomia, matematica, psicologia cognitiva, dell'organizzazione e sociale. Preece e colleghi (1994) invece ne propongono undici, differenziate in discipline che hanno dato un maggior contributo alla HCI (informatica, psicologia cognitiva, organizzativa e sociale, ergonomia e fattori umani) e discipline che sono interessate all'argomento (intelligenza artificiale, linguistica, filosofia, sociologia, antropologia, ingegneria e scienze della progettazione).

Spesso, però, la tanto acclamata interdisciplinarietà si risolve in una questione di facciata (Kaptelinin, 1993), sia per oggettive difficoltà legate alla comunicazione interdisciplinare, sia per la diffidenza con cui storicamente le singole discipline si sono confrontate. Al di là delle barriere istituzionali, il problema fondamentale è legato alla condivisione delle conoscenze (Matarazzo, 1994; Greece *et al.*, 1996). Le discipline che formano la HCI possiedono tutte apparati concettuali e metodologici chiaramente sviluppati, che rendono complessa la comprensione delle ricerche compiute nel loro ambito da parte di studiosi di altri settori. La continua proposta di contesti scientifici comuni, nella forma di congressi, libri o riviste, sembra comunque un segnale incoraggiante. Attualmente si è portati a credere che uno scambio proficuo possa nascere solo da un processo di mutuo accomodamento fra le discipline e non dall'imposizione di un unico stile ingegneristico come nel passato era stato proposto (§ 2.4). Questa idea è alla base delle più recenti proposte per l'insegnamento della HCI che prevedono corsi base di settori diversi (ACM SIGCHI, 1992; Preece *et al.* 1994). Difficilmente, secondo me, si giungerà alla piena integrazione delle conoscenze per cui uno stesso ricercatore possa essere in grado di progettare, valutare e realizzare sistemi interattivi. Penso piuttosto che la complessità del dominio richieda il lavoro comune di gruppi di ricercatori con specializzazioni diverse, ma con conoscenze fondamentali comuni.

2.3.1. Il contributo della psicologia

Il ruolo privilegiato che la psicologia dovrebbe giocare nella *human-computer interaction* è immediatamente sottolineato dal nome stesso della disciplina. Sfortunatamente però l'importanza nominale non si riflette a pieno nella realtà, dove il contributo psicologico, per quanto in sempre maggior espansione, è tuttora subordinato al ruolo delle scienze tecniche e frequentemente giudicato insoddisfacente (Newell e Card, 1985; Landauer, 1987; Green, Davies e Gilmore, 1996). Tale fenomeno può essere attribuito a due cause principali: (a) difficoltà di comunicazione con le scienze tecniche; (b) difficoltà di comunicazione con la psicologia accademica.

Tuttora molti tecnici faticano ad accettare che persone con competenze stereotipicamente umanistiche debbano partecipare alla progettazione (Coniglio e Paci, 1994). Inoltre, le scienze ingegneristiche e informatiche sono tradizionalmente considerate *più importanti*, in quanto specifiche depositarie di conoscenze relative alla progettazione, e *più affidabili*, in quanto scienze esatte. Nella HCI effettivamente il loro contributo viene *prima* di quello della psicologia, in quanto *prima* di iniziare qualunque analisi sul comportamento dell'utente è necessario verificare la fattibilità dello hardware e del software sottostanti l'interazione.

I rapporti fra HCI e psicologia accademica sono stati a lungo quasi inesistenti, precludendo così lo sviluppo di studi teorici e lavori empirici meno vincolati al puro lato applicativo²² (Green *et al.*, 1996). Spiegare questo generale atteggiamento di indifferenza può essere difficile. Certo, come vedremo più avanti (§ 3.4), gli esperimenti nel settore della HCI implicano un evidente allontanamento dal rigore sperimentale tipico della ricerca pura, ma questo oltre che un limite rappresenta una sfida. Piuttosto che rifiutare un settore in quanto difficilmente analizzabile, l'interesse psicologico si dovrebbe orientare verso la definizione di nuovi paradigmi esplorativi con cui aumentare il controllo sperimentale.

Esistono diversi modi in cui la psicologia può contribuire alla progettazione di sistemi informatici. Innanzitutto può fornire specifiche conoscenze teoriche sull'utente e un complesso apparato metodologico per l'analisi e lo studio del comportamento umano. In questo processo di semplice *trapianto teorico* sono state coinvolte la psicologia generale e, più recentemente, la psicologia sociale. La prima, che nella HCI è stata a lungo rappresentata dall'approccio cognitivista, può spiegare il funzionamento delle capacità mentali degli utenti. La seconda può aiutare a comprendere aspetti diversi del processo interattivo.

²² Questo problema è particolarmente forte in Italia, mentre negli Stati Uniti o in altri paesi Europei, quali Gran Bretagna e Olanda, numerose facoltà di Psicologia offrono corsi base e indirizzi di specializzazione in HCI.

La soluzione si è comunque rivelata insoddisfacente, con un'unica eccezione: l'applicazione al disegno di interfacce dei postulati di organizzazione percettiva della psicologia della gestalt (Borchers, Deussen e Knorzer, 1995). Invece, teorie e costrutti della psicologia cognitivista sono risultati inadeguati, perché tipicamente riguardano funzioni isolate e astratte dal loro contesto di azione (Landauer, 1987; Greece *et al.*, 1996). Il trapianto teorico della psicologia sociale appare ancora più problematico. Tale disciplina, infatti, si è sempre concentrata sullo studio del modo in cui gli individui attribuiscono un significato ed entrano in relazione con quei particolari oggetti della realtà che sono gli altri (Arcuri, 1985). La sostituzione di uno o più partner con dei computer implica inevitabilmente delle differenze di entità tale da precludere la generalizzabilità.

Anche il semplice trapianto metodologico implica notevoli difficoltà. Dato l'attuale stadio di sviluppo della HCI, ricerche effettivamente utili dovrebbero basarsi più su paradigmi esplorativi che su verifica di ipotesi (Landauer, 1987). I tipici lavori empirici sono solitamente ricerche sul campo o studi correlazionali, mentre il lato metodologico più robusto della psicologia si riferisce a esperimenti di laboratorio. Anche l'applicazione delle tecniche statistiche richiede profondi ripensamenti. Nel dominio in esame è importante evidenziare tendenze e definire la forza degli effetti, piuttosto che elencare livelli di probabilità (Landauer, 1987). Il settore statistico e metodologico richiedono alla psicologia il maggiore sforzo di adattamento linguistico per permettere la comprensione dei risultati anche ai non specialisti. Questo aspetto è stato fino ad ora prevalentemente ignorato generando da un lato lavori di ineccepibile qualità rimasti praticamente inutilizzati, dall'altro un generale disinteresse per l'aspetto metodologico (§ 3.4.1).

Affermare che la psicologia non è in grado di contribuire direttamente alla HCI non vuol dire negare la sua importanza. Anzi, quest'idea sottintende la necessità dello sviluppo di una specifica *psicologia della HCI*, settore finalizzato a produrre un apparato teorico e metodologico più consono al dominio di analisi. Ciò richiede innanzitutto un notevole impegno per creare adeguati strumenti di analisi, poi l'esecuzione di numerose ricerche empiriche per costruire modelli e teorie ed infine la traduzione delle teorie in un linguaggio facilmente comprensibile al settore tecnico. Quest'ultimo aspetto può svilupparsi su due direzioni: (a) traduzione esplicita, definizione di chiare direttive che il progettista deve apprendere; (b) traduzione implicita, incorporazione della teoria in strumenti di programmazione (Newell e Card, 1985; 1986).

Il valore di una specifica psicologia applicata alla HCI tende ad aumentare se si considera la natura biunivoca della relazione che la unisce alla psicologia pura. L'incontro fra le due aree di ricerca non si riduce all'innegabile contributo che teorie e metodi psicologici forniscono o hanno fornito alla psicologia applicata, ma include anche il contributo che studi relativi a questo particolare ambito del comportamento umano possono dare allo sviluppo e alla verifica di teorie prodotte in laboratorio (Newell e Card,

1986; Landauer, 1987). Conferme a questa asserzione si evidenziano fin dai primi studi di interazione utente-macchina. Non si dimentichi che proprio in tale contesto si svilupparono le idee che hanno portato all'affermazione del cognitivismo²³ (Luccio, 1982). Al di là di questo evento in sé eccezionale, si possono ricordare vari esempi in cui teorie dell'apprendimento e della memoria (Rieber, 1990; Mayer e Anderson, 1991, 1992; Mayer e Sims, 1994) o teorie di psicologia sociale (Spears, Lea e Lee, 1990; Lea e Spears, 1991, 1992; Spears e Lea, 1992, 1994) sono state sottoposte a verifica in situazioni di interazione utente-computer.

2.4. *Dal cognitivismo ai giorni nostri*

È opinione diffusa che le uscite pubbliche della psicologia dall'accademia avvengano in concomitanza con eventi eccezionali di grande rilievo, culturale, scientifico ed economico (Bagnara e Visciola, 1987). Seguendo questa chiave di lettura, l'evento determinante che ha spinto alcuni psicologi a uscire dai laboratori per collaborare con ingegneri e tecnici alla progettazione di macchine può essere fatto risalire alla seconda guerra mondiale. In quel periodo l'orientamento accademico prevalente era il comportamentismo, i cui famosi lavori sull'apprendimento vennero usati come modelli per l'ergonomia. All'inizio il comportamento di interazione utente-macchina venne dunque analizzato in puri termini di stimolo, risposta e rinforzo. L'assoluta mancanza di teorie sul funzionamento della mente umana produsse in generale lavori piuttosto sterili incapaci di descrivere in modo esaustivo il comportamento²⁴.

Il distacco da questa visione cominciò soprattutto in Europa, prima con i lavori di Craik e poi, in modo più sostanziale, con quelli di Broadbent. Dall'avvento del cognitivismo, lo studio dell'interazione utente-macchina fu prevalentemente caratterizzato da tale approccio. L'interesse verso il computer poi ha reso le relazioni fra cognitivismo e progettazione sempre più strette (Landauer, 1987). Infatti, il veloce sviluppo di nuove applicazioni informatiche rende il processo di apprendimento un fattore critico della prestazione; inoltre i compiti supportati dalle tecnologie informatiche coinvolgono in modo determinante le capacità della mente umana.

L'obiettivo principale della psicologia nella HCI divenne, dunque, la comprensione delle capacità cognitive sottostanti l'interazione (Preece *et al.*, 1994). Il cognitivismo deve

²³ In seguito all'analisi del comportamento di tracking, Craik (1943) fu il primo a definire l'utente come elaboratore di informazioni (Luccio, 1980). Nella sua visione l'elaboratore era identificato con un servo meccanismo di tipo cibernetico, al cui interno esisteva un unico meccanismo decisore caratterizzato da un funzionamento discreto. Egli ebbe inoltre l'intuizione di considerare il tempo di esecuzione come indicatore dei processi mentali sottostanti l'azione.

²⁴ Si consideri che attualmente il *carico cognitivo* viene considerato una delle variabili chiave della prestazione.

il predominio nella HCI al parallelismo fra il suo modo di concepire l'essere umano e l'effettiva struttura del computer. Proprio tale similitudine ha facilitato l'incontro e lo scambio fra informatici e psicologi. Secondo l'approccio cognitivista l'interazione può essere schematicamente rappresentata da un ciclo di elaborazione dell'informazione composto da due unità di elaborazione, l'essere umano e il computer, che, pur essendo strutturalmente e cognitivamente differenti, possono essere descritte in modo simile (vedi Figura 1).

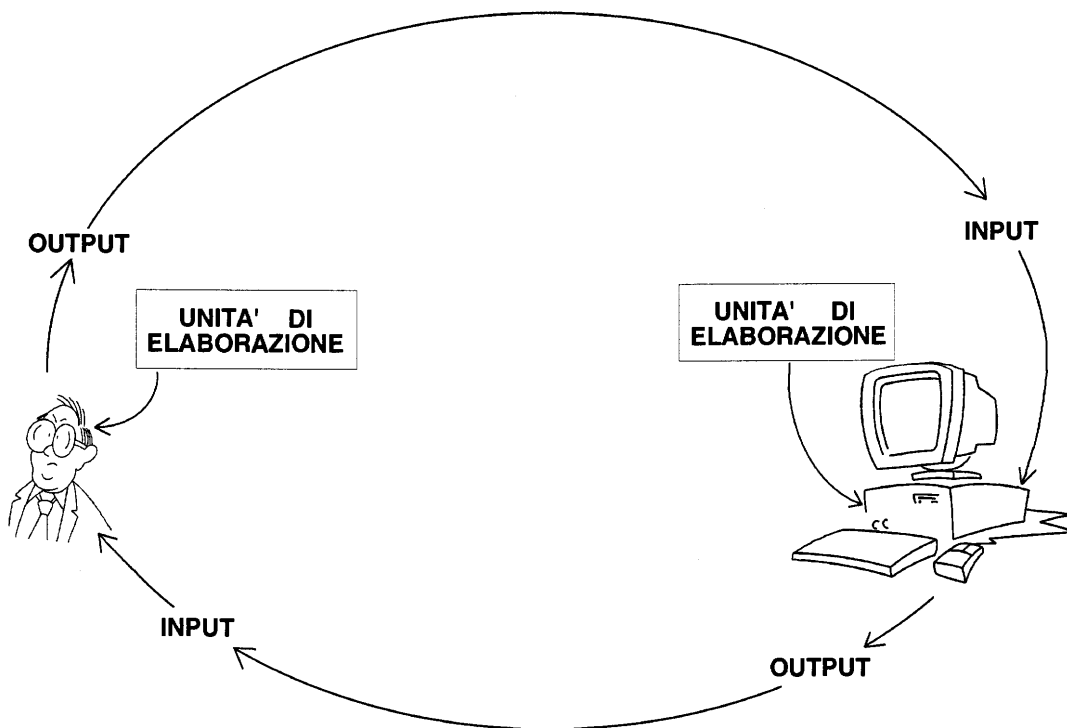


Figura 1. Ciclo di elaborazione dell'informazione secondo l'approccio cognitivista.

L'output di un'unità rappresenta l'input per l'altra e l'interazione si sviluppa attraverso continui cicli di elaborazione prodotti dalle singole unità.

Il primo contributo programmatico della psicologia cognitiva allo studio dell'interazione utente-computer risale alle ricerche iniziate nel 1974 da un gruppo di ricercatori dello XEROX PARC (Palo Alto Research Center). Dal momento che la loro teoria ha esercitato un'influenza fondamentale, per almeno dieci anni, nel determinare scopi e metodi della psicologia nella HCI proviamo a descriverla brevemente.

2.4.1. La psicologia ingegneristica

Attraverso l'esecuzione di un'impressionante quantità di studi sperimentali (per una dettagliata bibliografia si veda Card, Moran e Newell, 1983) i ricercatori dello XEROX Parc giunsero alla definizione di una *psicologia cognitiva applicata all'interazione utente-computer* (Card et al., 1983; Newell e Card, 1985; 1986). Il loro approccio era caratterizzato da un prevalente stile ingegneristico, considerato necessario alla comunicazione interdisciplinare dei risultati. Secondo gli autori, infatti, solo tale caratteristica poteva garantire alla psicologia l'importanza che le spetterebbe di diritto nella HCI ed evitare il posto di subordine cui era stata relegata l'ergonomia (Newell e Card, 1985; 1986). A tale fine era stato necessario ridefinire il dominio della psicologia nel contesto della HCI, restringendolo alle sole azioni caratterizzate da un tempo di esecuzione compreso fra il decimo di secondo e il minuto. In tale fascia si trovano processi simbolici e meccanismi mentali. Le attività più lunghe, guidate da obiettivi e scopi, venivano considerate dominio della razionalità limitata e quindi non di pertinenza della psicologia.

La psicologia ingegneristica ha fornito due contributi fondamentali: (a) il Modello dell'Elaboratore Umano, che rappresenta un mirabile, seppur limitato, esempio di adeguamento di teorie cognitiviste al dominio della HCI (vedi Figura 2); (b) una famiglia di modelli di analisi del compito indicati dall'acronimo GOMS (§ 3.3).

La psicologia ingegneristica si basa su tre concetti fondamentali: *analisi del compito, operazioni e approssimazione*. L'obiettivo principale è predire il comportamento dell'utente partendo da una dettagliata descrizione simbolica del compito che egli deve eseguire. La formalizzazione viene eseguita tramite analisi GOMS e si sviluppa mediante esplicite operazioni matematiche, le cui costanti vengono inferite dal Modello del Processore Umano. Seguendo tale procedura si ottengono dei modelli cognitivi *approssimati* dell'utente come elaboratore di informazioni. Il termine approssimato indica che le predizioni del modello presentano sempre un margine di errore dovuto alla difficoltà di prevedere e comprendere i comportamenti erranei. Postulato fondamentale della teoria è la sequenzialità temporale delle azioni analizzate. Per questo, e per altri limiti di cui discuteremo più avanti, tali modelli si riferiscono prevalentemente al comportamento di utenti esperti nella digitazione di lettere o comandi.

Secondo questa prospettiva, scopo principale della psicologia nella HCI è estendere e raccogliere teorie generali piuttosto che valutare particolari progetti (Newell e Card, 1985; 1986). Ciò dipende dall'assunto che il risultato del processo di progettazione è vincolato fin dal primo momento della sua stesura e non dipende da successivi momenti di valutazione.

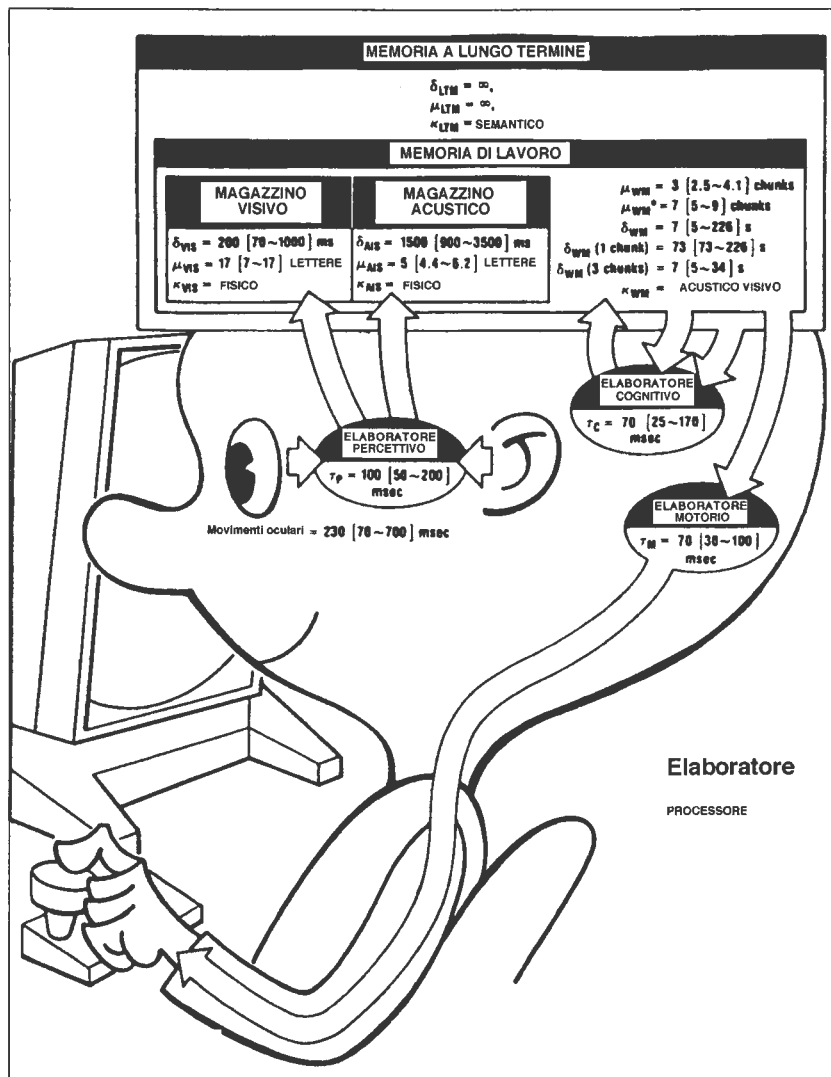


Figura 2. Modello dell'Elaboratore Umano (Adattato da Card *et al.*, 1983, pag. 26).

Attraverso l'Elaboratore Percettivo l'informazione sensoria giunge alla memoria di lavoro, composta da *chunk* attivi della memoria di lavoro. Il principio di funzionamento dell'Elaboratore Umano è il ciclo Riconoscimento-Azione diretto dall'Elaboratore Cognitivo. L'Elaboratore Motorio è attivato dalla memoria di lavoro.

2.4.2. Oltre la psicologia ingegneristica

L'importanza della psicologia ingegneristica può essere testimoniata, oltre che dalle numerose ricerche che a essa si ispirano, anche dal numero di critiche che fin dai suoi albori ha generato. (Carroll e Campbell, 1986; Karat, 1988). A tale proposito è interessante notare che il più approfondito elenco dei limiti sottostanti la teoria viene fornito dagli autori stessi nella prima presentazione del modello GOMS (Card, Newell e Moran, 1980).

Le più severe critiche all'approccio riduzionista nacquero in seno alla stessa corrente cognitivista. Ne è un segno la polemica che, sulle pagine della rivista *Human-Computer Interaction*, contrappose i sostenitori della psicologia ingegneristica (Newell e Card, 1985; 1986) a Carroll e Campbell (1986). Questi ultimi criticavano innanzitutto l'idea della trasformazione della psicologia in scienza ingegneristica, non solo a causa dell'inadeguatezza del dominio ma anche a causa dell'inadeguatezza della metodologia (Carroll e Campbell, 1986). Ridurre la psicologia al solo studio di processi simbolici e meccanismi mentali implica lo snaturamento della disciplina stessa. Inoltre, l'approccio ingegneristico privilegia la produzione di modelli descrittivi e tende a ignorare i modelli esplicativi utili in psicologia²⁵.

All'idea di un radicale cambiamento della psicologia viene contrapposta l'esigenza di una piena interdisciplinarietà, sorretta dal contemporaneo coinvolgimento delle varie aree di studio nella progettazione e valutazione di prototipi. Ciò implica una ridefinizione degli scopi della psicologia (Carroll e Campbell, 1986; Lewis, 1990): dalla definizione di teorie generali sul funzionamento cognitivo dell'essere umano si passa alla ricerca di metodologie sperimentali per valutare come i sistemi informatici vengono effettivamente utilizzati. L'approccio cognitivista viene così integrato dalla corrente ecologica, la cui importanza si fa sentire nell'esigenza di considerare il contesto in cui l'informazione viene elaborata (Norman e Draper, 1985, Woods e Roth, 1989; Lewis, 1990; Norman, 1991).

Alla base della crisi della psicologia ingegneristica può forse essere situato il crescente interesse che, verso la metà degli anni ottanta, ha raccolto lo studio dell'errore umano, considerato come un evento naturale nel comportamento di interazione (Lewis e Norman, 1986; Rizzo, Bagnara e Visciola, M., 1987; Norman, 1988). L'incompatibilità fra psicologia ingegneristica e studio dell'errore è evidente, sia per il dominio in cui esso deve venire studiato (azioni finalizzate a uno scopo), sia per l'inadeguatezza dell'apparato metodologico appropriato solo alla descrizione di prestazioni perfette (§ 3.3). L'estensione del dominio psicologico ha come conseguenza un crescente interesse verso il compito (Carroll *et al.*, 1991) e verso l'analisi delle modalità di rappresentazione della conoscenza, per esempio i modelli mentali con cui gli utenti si rappresentano il funzionamento del sistema (Carroll e Olson, 1988) e le metafore con cui comunicare all'utente il funzionamento del sistema (§ 3.2.1).

A partire dalla fine degli anni '80, il cognitivismo riduzionistico di Card e colleghi appare definitivamente respinto. In contrapposizione si fa strada l'idea che l'apporto

²⁵ Tale distinzione è fondamentale (Campbell e Bickhard, 1986). Un modello esplicativo cerca di spiegare i processi mentali attraverso i quali l'utente esegue un compito. Un modello descrittivo, invece, predice le prestazioni nell'esecuzione del compito, senza spiegare i mezzi attraverso i quali l'utente li esegue. I primi spiegano dunque *perché* certi fenomeni avvengono, i secondi *come* essi avvengono. Nella tradizione ingegneristica la descrizione è subordinata all'azione, al controllo e alla possibilità di modifica dei fenomeni descritti (Re, 1995).

psicologico alla HCI non possa limitarsi solo all'analisi cognitiva di processi elementari, ma richieda il coinvolgimento di più settori della psicologia (Carroll, 1987; Norman, 1987; Carroll *et al.*, 1991). Negli ultimi anni si alternano tentativi di ripensamento interni alla corrente cognitivista e posizioni di rifiuto radicale di tale approccio. I primi propongono essenzialmente una riduzione dell'approccio costruttivista e mentalista nonché una maggiore attenzione all'*utente che agisce in un contesto* (Norman e Draper, 1985; Landauer, 1987; Lewis, 1990). Secondo Bannon (1991) ciò implica un chiaro mutamento nella definizione di utente da *fattore umano*, tradizionalmente inteso come agente passivo e spersonalizzato, ad *attore umano*, agente attivo capace di controllo e di scelta.

Contemporaneamente si assiste alla proposta e alla successiva affermazione di paradigmi di ricerca alternativi, quali la *teoria dell'attività* (Bodker, 1989; Kaptelinin, 1993) e a varie proposte derivate dalla psicologia sociale che sanciscono la sterilità di un approccio che non consideri il contesto in cui l'azione viene eseguita. L'estensione delle conoscenze sociali al dominio della HCI dovrebbe procedere nel seguente modo (Naas *et al.*, 1994):

1. Determinazione di un particolare effetto o costrutto sociale da sottoporre a verifica.
2. Sostituzione di uno o più umani con computer.
3. Verifica empirica della nuova soluzione.
4. Confronto delle due situazioni.

In generale si assiste dunque a uno spostamento di interesse: dalla comprensione delle capacità cognitive dell'utente allo studio del processo interattivo (Lewis, 1990; Preece *et al.* 1994).

2.5. Conclusioni

Nel corso della sua breve storia la HCI si è andata velocemente evolvendo. Il suo sviluppo può essere categorizzato in tre stadi principali:

- *Prospettiva Sistemica.* La HCI è concepita come analisi del rapporto fra utenti e computer: attore dell'interazione è dunque il sistema utente-artefatto. La psicologia si concentra sullo studio del sistema di elaborazione umano, allo scopo di produrre teorie generali del suo funzionamento durante l'interazione. Il cognitivismo rappresenta il paradigma teorico principale.
- *Prospettiva Personale.* L'attenzione si rivolge anche al compito nella convinzione che l'utente interagisca col computer al fine di eseguire un compito che a causa dell'artefatto tende a cambiare la sua natura. La psicologia si concentra sullo studio dei processi mentali superiori coinvolti nell'interazione, allo scopo di valutare la prestazione umana durante l'interazione con particolari prototipi. Il cognitivismo viene affiancato dalla corrente ecologica.
- *Prospettiva interazionale.* L'interesse si allarga fino a includere l'ambiente in cui utenti e computer eseguono il compito. La psicologia si concentra sullo studio delle caratteristiche dell'interazione. Cognitivismo e corrente ecologica sono affiancati dalla psicologia sociale.

3. Misure e metodologie di valutazione

La qualità di un sistema informatico è solitamente definita come amichevolezza (*friendliness*) o, più rigorosamente, come usabilità (*usability*). Tale concetto sta assumendo un ruolo sempre più importante nello sviluppo dei nuovi sistemi informatici. Negli ultimi 20 anni il budget preposto per le valutazioni di usabilità è raddoppiato (Nielsen, 1994) e le più importanti compagnie informatiche si sono dotate di esperti, laboratori e sofisticate attrezzature per la misurazione della qualità dei sistemi (Matarazzo, 1994; Coniglio e Paci, 1994).

L'esigenza di sviluppare sistemi altamente usabili deriva dall'azione congiunta di tre fattori principali.

- *Sviluppo tecnologico.* La rapida evoluzione di hardware e software permette oggi macchine capaci di eseguire compiti impensabili in passato (Myer, 1996). Fra i tanti esempi si pensi ai personal computer, al mouse, alle interfacce grafiche o allo sviluppo delle reti elettroniche.
- *Motivazioni economiche.* Nei paesi industrializzati l'informatica è diventata una delle maggiori fonti di ricchezza, sia perché produce denaro sia perché lo gestisce. Consideriamo il primo aspetto: l'industria si è rivolta all'approccio *centrato sull'utente* per aprirsi nuovi mercati. La rivoluzione dei personal computer, avvenuta in un regime di acerrima concorrenza, ha prodotto un netto decremento nei prezzi di hardware e software, rendendo il calcolatore disponibile a un'utenza sempre più vasta. Col tempo il mercato convenzionale si è andato saturando e di conseguenza è diventato più esigente. L'unica via di ripresa consiste nel trovare *nuovi usi, nuove modalità d'uso, nuovi utenti*. Dal punto di vista della gestione del denaro, l'esigenza di usabilità invece deriva dalla larga applicazione di tecnologie informatiche in ambienti centrali per lo sviluppo economico. Una buona qualità nell'interazione è necessaria a garantire sicurezza nella gestione e diminuzione degli sprechi. Il problema della sicurezza verrà affrontato fra breve, una dettagliata esposizione di esempi di risparmio indotto dalla realizzazione di sistemi adeguati alle capacità umane è riportata da Nielsen (1993).
- *Esigenza di sicurezza.* Il crescente controllo che i computer hanno su attività vitali per il genere umano ha subito posto il problema della sicurezza. Il settore in cui tale fattore ha

avuto un'importanza fondamentale nel graduale processo di semplificazione dell'interazione, è l'ambiente aeronautico (Winter, 1986). Negli ultimi cinquant'anni gli aeroplani si sono andati evolvendo da semplici macchine guidate manualmente fino ai complessi e sofisticati jet del giorno d'oggi. In questo periodo, lo sviluppo tecnologico ha completamente modificato la natura degli errori che possono scatenare incidenti. Si evidenzia un netto decremento di errori dovuti alla macchina e un parallelo aumento di errori attribuibili all'operatore. Anche la natura della responsabilità umana ha subito un importante cambiamento. Negli aerei meno evoluti gli errori umani erano riconducibili a manovre errate o a sbagli nella lettura degli strumenti di bordo. Nei moderni aerei, invece, gli errori dipendono in massima parte da situazioni interattive fallimentari fra l'operatore e le sofisticate strumentazioni. Essi consistono in errori nella programmazione, nell'immissione di informazioni o nell'interpretazione degli output in situazioni fuori dall'ordinario. È chiaro che in questo contesto l'usabilità diventa un fattore cruciale della progettazione.

L'usabilità di un sistema può essere misurata tramite approcci profondamente diversi. Secondo la tassonomia proposta da Karat (1989) esistono tre metodologie principali: valutazione basata sulla teoria (§ 3.2), valutazione basata sul compito (§ 3.3), valutazione basata sugli utenti (§3.4). Ognuna di esse è caratterizzata da pregi e difetti, relativi a costi, validità e generalizzabilità dei risultati, che verranno sottolineati nel corso dell'esposizione.

3.1. Usabilità dei sistemi

L'usabilità è un complesso concetto di natura multidimensionale, caratterizzato in modo variabile a seconda dei modelli; Senza la pretesa di essere esaustivi, ci soffermeremo su due importanti definizioni: quella elaborata da Nielsen (1993) e quella fornita dagli ISO/CD 9241-11 (International Standard Organization/Computer Display).

La prima ha il pregio di presentare un dettagliato modello degli attributi costituenti l'usabilità e delle sue relazioni gerarchiche con altri concetti rilevanti per il dominio. La seconda, invece, rappresenta uno standard per la comunità degli ergonomi (Beimel, Schindler, Wandke, 1994; Harker, 1995 ; Stewart, 1995).

3.1.1. La proposta di Nielsen

Secondo il modello gerarchico di Nielsen, l'usabilità è un aspetto dell'accettabilità di un sistema vale a dire della sua capacità globale di soddisfare necessità e richieste degli utenti. Tale concetto si articola in due sottocategorie: *accettabilità sociale* e *accettabilità pratica*.

Da un punto di vista sociale un sistema viene accettato se ritenuto coerente con alcuni valori fondamentali dell'essere umano, quali, per esempio, sicurezza e privacy. Ovviamente il giudizio dipende in maniera cruciale dalla cultura dell'utente e dal contesto di utilizzo. Da un punto di vista pratico l'accettabilità è invece definita da una serie di attributi quali: costo, affidabilità, compatibilità con sistemi preesistenti e praticità. Quest'ultimo aspetto a sua volta è definito da due attributi: *utilità* e *usabilità*. Il primo si riferisce all'effettiva capacità del sistema di eseguire le funzioni per cui è stato programmato, mentre il secondo considera la qualità dell'interazione richiesta per l'esecuzione di tali funzioni.

L'usabilità è dunque il concetto chiave nel rapporto tra utente e computer. Essa è composta da cinque attributi basilari, oggettivamente ed empiricamente valutabili.

- *Facilità di apprendimento*. Un utente, che utilizza per la prima volta un sistema, deve raggiungere buone prestazioni in tempi brevi. Per definire la priorità di tale attributo è necessario prevedere il grado di esperienza dell'utenza target e la durata media dell'interazione. Solo così è possibile ottenere un buon compromesso fra facilità di apprendimento e efficienza del sistema.
- *Efficienza nell'utilizzo*. Una volta superata la fase di apprendimento il sistema deve garantire prestazioni altamente produttive. Questo attributo viene solitamente misurato con utenti caratterizzati da un discreto e stabile livello di esperienza.
- *Facilità di ricordo*. L'utente deve potere interagire con un'interfaccia anche dopo un periodo di lungo inutilizzo, senza essere costretto ogni volta a ricominciare da zero il periodo di apprendimento. Una semplice e immediata memorizzazione dei comandi è particolarmente importante per applicazioni destinate ad un utilizzo occasionale. Una strategia standard per migliorare il ricordo di specifici comandi è il riconoscimento supportato dall'utilizzo di icone o di metafore (§ 3.2.1). Un'alternativa interessante, ma attualmente poco studiata, è l'uso di procedure *bizzarre*. Tale soluzione sfrutta una caratteristica della memoria umana per cui, quando si utilizza l'immaginazione per aiutare il ricordo è preferibile costruirsi delle immagini bizzarre che, colpendo e

attraendo maggiormente l'attenzione, favoriscono la ritenzione permanente di informazioni nuove²⁶ (Delin, 1968). La soluzione comunque va usata con estrema cautela per non disorientare eccessivamente l'utente.

- *Sicurezza e robustezza all'errore.* Il sistema deve essere il più possibile sicuro, deve cioè ridurre al minimo la probabilità del verificarsi di azioni errate. Inoltre deve essere robusto, nel senso che il suo funzionamento non deve venire compromesso da errori banali e deve permettere semplici azioni riparatorie.
- *Soddisfazione soggettiva.* Questa è l'unica dimensione soggettiva e riguarda il giudizio relativo al comfort e alla piacevolezza dell'interazione. In generale, gli utenti dovrebbero sentirsi soddisfatti del tempo passato utilizzando il computer.

Il valore relativo dei cinque attributi è variabile. Infatti, vari vincoli inerenti alla progettazione e realizzazione non consentono sempre di ottenere valori elevati per ciascuno di essi. È quindi necessario ordinare gli attributi, sulla base delle caratteristiche dell'utenza e del sistema, su una scala di importanza. Si consideri per esempio la differenza fra un'applicazione per esperti, che può richiedere un maggiore sforzo di apprendimento ma deve consentire prestazioni elevate, con un'applicazione per inesperti, che deve richiedere un minimo sforzo di utilizzo anche a scapito dell'efficienza.

3.1.2. La proposta ISO/CD 9241

Secondo gli standard ISO/CD 9241 l'usabilità è definita come l'*efficacia*, *efficienza* e *soddisfazione* con cui determinati utenti raggiungono determinati scopi in un ambiente ben definito. La dimensione dell'*efficacia* riguarda il grado di accuratezza e completezza con cui gli utenti raggiungono lo scopo. La dimensione dell'*efficienza* misura il rapporto fra efficacia e risorse spese per il raggiungimento dello scopo. La *soddisfazione dell'utente* è concettualizzata in modo simile alla proposta di Nielsen.

Per misurare l'usabilità di un sistema è necessario operationalizzare le tre dimensioni definendo, per ognuna di esse, gli attributi fondamentali sulla base del sistema oggetto di valutazione e del suo *contesto d'uso*. In questa visione il termine contesto si riferisce a tutti gli attributi rilevanti ai fini dell'usabilità, che riguardano alcune caratteristiche degli utenti, del compito, dello hardware e del software come anche

²⁶ Un buon esempio di procedura bizzarra è costituito dall'estrazione del dischetto in ambiente Macintosh mediante trascinalamento nel cestino. Tale azione si pone in contrasto con buona parte dei principi di progettazione. È infatti, incongruente e concettualmente incompatibile (Sanders e McCornick, 1992) con la funzione base del cestino, cioè la cancellazione dal disco rigido degli elementi introdotti. Eppure la procedura di estrazione, per quanto inizialmente traumatica, difficilmente verrà dimenticata, anche dopo un lungo periodo di inutilizzo.

dell'ambiente in cui il sistema viene utilizzato. Efficacia ed efficienza sono misurate da indici *oggettivi* di prestazione, per esempio numero di errori nel risultato finale (*efficacia*), o numero di correzioni (*efficienza*). Tali indici possono essere ottenuti da valutazioni basate sulla teoria e mediante sperimentazione empirica, per quanto le ultime solitamente forniscono risultati più attendibili. La soddisfazione dell'utente invece è misurata da indici *soggettivi*, raccolti tramite questionari o interviste strutturate. I giudizi spaziano da valutazioni di carattere prettamente utilitaristico fino a valutazioni edonistiche e di desiderabilità del sistema.

3.1.3. Usability Engineering

L'importanza della fase di valutazione è testimoniata dal recente emergere di una specifica area di studio dell'ingegneria, denominata *Usability Engineering* (Gould e Lewis, 1985; Whiteside, Bennet e Holtzblatt, 1988; Nielsen, 1993). In quanto disciplina ingegneristica il suo obiettivo generale consiste nella progettazione e costruzione di artefatti con scopi pratici (Whiteside et al., 1988). Il suo obiettivo specifico è invece l'elaborazione di metodologie di progettazione che guidino la realizzazione di sistemi altamente usabili.

Seguendo uno dei canoni fondamentali dell'ergonomia cognitiva, la *Usability Engineering* suggerisce che la valutazione non deve rappresentare lo stadio finale del tradizionale processo di progettazione *top-down* (Nielsen, 1993). Tale metodologia, basata su un approccio centrato sul sistema, prevede una procedura strettamente sequenziale che comincia con la completa specificazione del sistema sulla base del problema da risolvere. Il sistema viene quindi implementato, attraverso varie fasi di miglioramento (*stepwise refinement*). Il limite di tale approccio consiste però nell'eventualità che nel corso della progettazione una delle caratteristiche iniziali debba essere modificata. Tale possibilità diventa tanto più frequente quanto più il sistema è complesso e il fattore umano, con le sue richieste e i suoi personali desideri, diventa parte della progettazione.

Secondo la metodologia proposta dalla *Usability Engineering* l'usabilità deve essere considerata fin dall'inizio della progettazione (vedi Figura 3). Ciò implica che, prima ancora di definire le strategie di soluzione del sistema, si determini il livello di usabilità desiderato. A tale fine è necessario definire accuratamente gli attributi di usabilità salienti, operazionalizzandoli e ordinandoli su una scala di importanza, nonché determinare per ognuno di essi un livello minimo (Good, Spine, Whiteside e George, 1986). Il livello è misura dell'importanza assegnata al fattore umano e deriva da attente mediazioni fra tale aspetto, fattori economici e limiti tecnologici.

Il sistema viene poi sviluppato tramite varie tecniche. Una fra le più utili è la seguente. Si parte dalla progettazione parallela (*parallel design*) di più prototipi. Le diverse soluzioni vengono poi valutate empiricamente e la migliore viene ulteriormente elaborata tramite progettazione iterativa (*iterative design*; Whiteside, 1988; Gould, 1988; Shneiderman, 1992) fino al raggiungimento del criterio di usabilità stabilito. La progettazione iterativa implica una sorta di *loop* fra progettazione, valutazione e implementazione. Il parallelo evolvere di specificazione e implementazione permette di evitare cambiamenti radicali in prodotti già pienamente sviluppati e quindi difficilmente modificabili. Ovviamente l'intera procedura sottintende l'utilizzo di strumenti efficaci per una veloce realizzazione e validazione di prototipi.

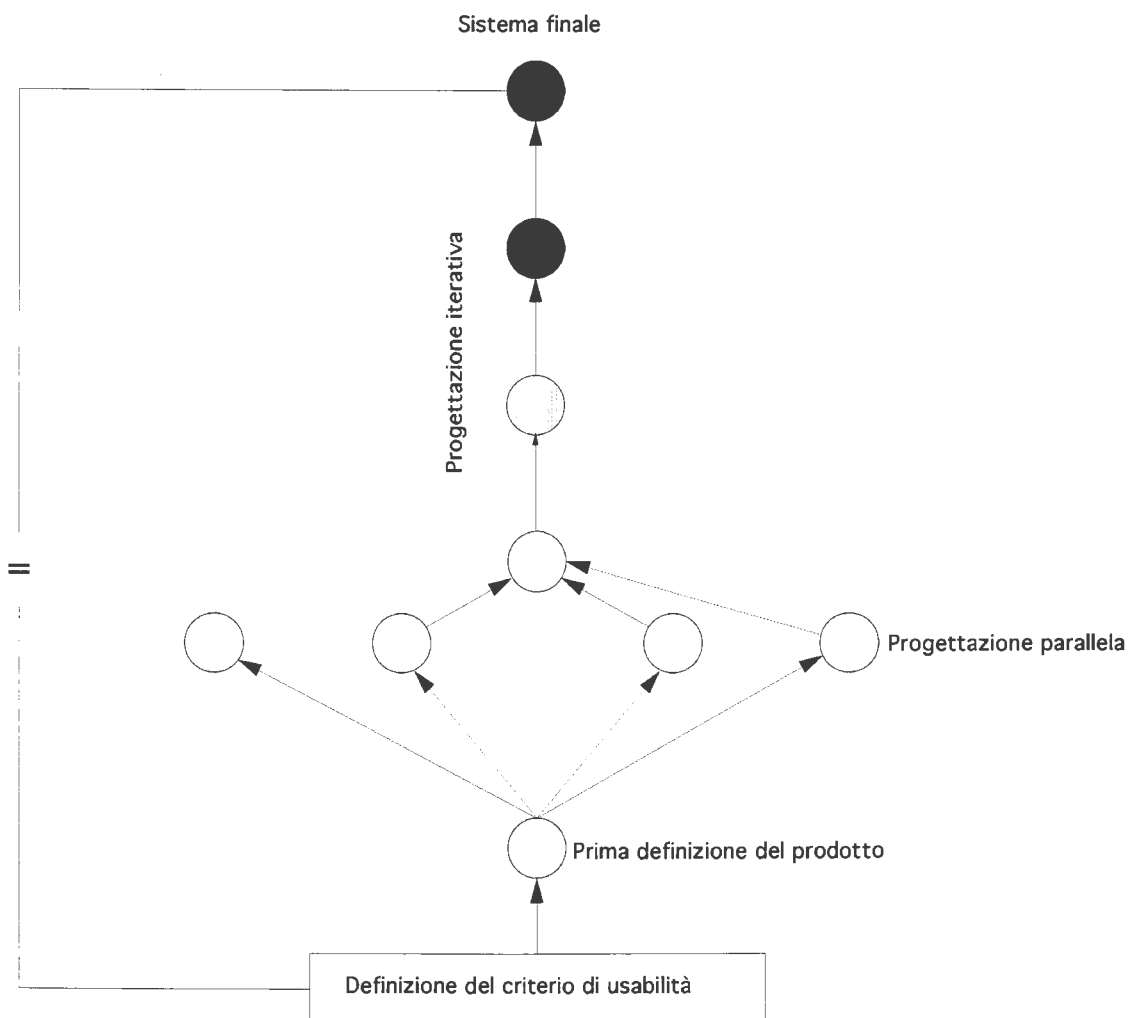


Figura 3. Esempio di progettazione secondo i criteri stabiliti dalla Usability Engineering.

Si noti che sotto molti aspetti la *Usability Engineering* rappresenta un mirabile esempio di integrazione di tecniche della psicologia sperimentale all'interno del processo di progettazione (§ 2.3.1). In essa il contributo psicologico, nella forma di valutazione

empirica e teorica della prestazione, tende ad assumere pari importanza di quello informatico. Peccato però che la nuova proposta si stia sviluppando prevalentemente in ambito tecnico, sottovalutando e sottovalutato dalla psicologia.

3.2. *Valutazione basata sulla teoria*

Le valutazioni basate sulla teoria sono valutazioni di tipo euristico in cui viene verificato l'adeguamento delle caratteristiche di un sistema a un elenco di direttive di usabilità (Nielsen e Molich, 1990). Tipicamente tale processo richiede il coinvolgimento di più valutatori esperti²⁷.

Il mondo informatico da sempre ha riposto grandi aspettative verso questa forma di valutazione. Tutti i più importanti testi sulle interfacce dedicano ampio spazio a elenchi di direttive di progettazione, spesso differenziate a seconda della loro generalizzabilità (Shneiderman, 1992; Timpleby, 1980). Attualmente disponiamo di numerose liste di *direttive generali*²⁸ (regole base applicabili a tutti i tipi di interfacce), alcune di *direttive categoriali* (regole relative a specifici sistemi, per esempio direttive per sistemi di video scrittura), ma estremamente pochi esempi di *direttive specifiche* (regole relative alle caratteristiche di un particolare prodotto). Un famoso esempio di direttive generali è costituito dalle *otto regole d'oro* elaborate da Shneiderman (1992; vedi Finestra 2). Si può dunque sostenere che le valutazioni basate sulla teoria sono potenzialmente molto utili, ma la teoria disponibile è ancora troppo vaga a garantire un loro effettivo successo. In futuro è auspicabile che questa metodologia diventi la tecnica di verifica principale. Infatti non solo permette analisi a basso costo e relativamente veloci, ma rappresenta anche un potente strumento su cui i progettisti possono direttamente basare il loro lavoro. Nel frattempo la valutazione teorica è particolarmente utile per generare dei modelli teorici di massimizzazione dell'usabilità piuttosto che per effettive verifiche del loro successo pratico.

²⁷ Paragonando i risultati di sei diverse valutazioni euristiche, Nielsen (1992) ha dimostrato che in media un solo esperto è in grado di individuare appena il 35% dei problemi di usabilità. Il numero ideale di valutatori secondo un'analisi costi-benefici è di circa cinque persone.

²⁸ Fra le raccolte più estese ricordiamo: Smith e Mosier (1986) con 944 direttive; Marshall et. al. (1987) con 162 direttive; Brown (1988) con 302 direttive; Mayhew (1992) con 288 direttive (riferimenti bibliografici in Nielsen, 1993).

Un sistema usabile deve:

1. Essere consistente: scopi uguali devono sempre essere raggiungibili da procedure uguali
2. Permettere l'uso di tasti funzioni (§)
3. Fornire feedback informativi dopo ogni azione dell'utente
4. Progettare dialoghi mirati alla chiusura: le sequenze di azioni devono essere organizzate in strutture caratterizzate da un inizio, un punto di mezzo e una fine.
5. Permettere un agevole recupero dagli errori
6. Permettere un semplice annullamento dell'azione
7. Sostenere un atteggiamento di *locus of control* interno: l'utente deve sentire di avere in mano il controllo dell'interazione.
8. Diminuire il carico della memoria a breve termine, rispettando i limiti mnemonici del sistema cognitivo umano.

Finestra 2. Le otto regole d'oro per un'interfaccia usabile (Shneiderman, 1992)

3.2.1. Modelli teorici di massimizzazione dell'usabilità

Attualmente sono disponibili più modelli di massimizzazione dell'usabilità, che essenzialmente fanno riferimento a due diverse modalità di funzionamento cognitivo: l'inferenza e l'analogia.

Nel modello basato sull'inferenza il principio di minimo svolge un ruolo cruciale: l'interazione è valutata in funzione del numero di azioni da eseguire, nonché del numero di concetti e procedure da ricordare (Card, Moran e Newell, 1983; Polson e Kieras, 1985; Archer e Yuan, 1995). L'usabilità è quindi definita in base alla *complessità del sistema* che risulta minore quando viene ridotto il numero di azioni richieste all'operatore. Applicazioni pratiche di questa idea sono per esempio i tasti funzione e alcune forme di menu (§ 1.2).

Il modello analogico è molto noto, grazie allo straordinario successo delle interfacce grafiche (§1.2). La filosofia sottostante sfrutta la similitudine con conoscenze precedenti come mezzo per rendere l'interazione più efficiente e immediata. La complessità di concetti, azioni e processi informatici viene diminuita mediante metafore che creano

un'associazione fra il dominio informatico e un altro dominio più familiare all'utente (Carroll, Mack e Kellogg, 1988; Marcus, 1993).

È auspicabile che in futuro vengano sviluppati dei modelli basati su teorie sociali per la massimizzazione dell'usabilità di sistemi flessibili. Tuttavia, questo appare un compito estremamente delicato in cui va sempre tenuta presente la peculiarità del dominio (§2.1.1).

3.3. Valutazione basata sull'analisi del compito

La valutazione mediante analisi del compito si pone a distanza intermedia fra le altre due metodologie. La prestazione viene valutata dall'analisi formale delle procedure di esecuzione in riferimento a una concettualizzazione astratta dell'utente, conosciuta come utente *ideale*. L'analisi del compito si sviluppa secondo un procedimento induttivo, partendo cioè dalla definizione del compito a un alto livello di astrazione e scendendo via via fino all'identificazione delle azioni elementari. La formalizzazione permette di identificare eventuali difetti delle strategie di soluzione previste (Nielsen, 1993).

Esistono svariati metodi di analisi del compito. Sulla base dell'approccio privilegiato essi possono essere classificati in tre categorie:

- metodi derivati dall'ingegneria del software
- metodi derivati dall'intelligenza artificiale
- metodi di tipo cognitivo

Alcuni esempi di metodi derivati dall'ingegneria del software sono la tecnica di analisi e progettazione strutturata (*Structured Analysis and Design Technique*, SADT), le tavole decisionali e i diagrammi di passaggio di stato. Dall'Intelligenza Artificiale derivano invece le pianificazioni gerarchiche e le reti semantiche. Attualmente però i metodi di tipo cognitivo sono considerati i più adatti per la fase di valutazione.

3.3.1. Metodi di tipo cognitivo

Fra i metodi di tipo cognitivo il più noto è l'analisi GOMS (§ 2.4) che rappresenta una famiglia gerarchica di modelli. Le lettere dell'acronimo indicano gli elementi costitutivi della conoscenza necessaria all'utente per eseguire un compito: *Goals* (obiettivi), *Operators* (operatori), *Methods* (metodi) e *Selection Rules* (regole di selezione). Il comportamento dell'utente viene descritto partendo dallo scopo superiore sottostante la sua azione (*top-level goal*) e poi analizzato mediante una strategia di suddivisione in sotto-scopi (*sub-goals*), a loro volta articolati in scopi sempre più specifici.

In questa visione uno scopo è una struttura simbolica che definisce lo stato meta e che determina un insieme di possibili metodi necessari al suo raggiungimento. Gli *operatori* di base sono invece atti fisici, percettivi o cognitivi la cui esecuzione produce un cambiamento nello stato del compito o nella stato mentale dell'utente. Un *metodo* è una sequenza di operatori necessari al raggiungimento di uno scopo. A volte metodi diversi possono portare al raggiungimento dello stesso scopo. In questi casi è dunque necessario specificare le *regole di selezione*, cioè le condizioni sulla cui base un utente può scegliere fra i vari metodi disponibili. Un esempio di analisi GOMS è riportato nella Finestra Finestra 3.

```

GOAL: attiva icona x                               top level goal
. [select Goal: attiva MEDIANTE DUE AZIONI DISTINTE   if utente = inesperto
. . . GOAL: SELEZIONA ICONA
. . . . SCEGLI COMANDO
. . . . . [select: GOAL: SELEZIONA TRAMITE MOUSE
. . . . . . CLICK SU ICONA X
. . . . . . GOAL: SELEZIONA TRAMITE TASTIERA
. . . . . . . DIGITA LETTERA O TASTO SPECIALE]
. . . . . GOAL: CONFERMA SELEZIONE
. . . . . . SCEGLI COMANDO
. . . . . . . [select GOAL : CONFERMA TRAMITE MOUSE
. . . . . . . . CLICK SU ICONA OK
. . . . . . . GOAL: CONFERMA TRAMITE TASTIERA
. . . . . . . . ENTER]
. GOAL: ATTIVA ICONA X CON UN'UNICA AZIONE scorciatoia per utenti esperti
. . SCEGLI COMANDO
. . . [select: GOAL: ATTIVA DA TASTIERA
. . . . DIGITA LETTERA+ TASTO SPECIALE
. . . . GOAL: ATTIVA DA MOUSE
. . . . . DOPPIO CLICK SU ICONA] ]

```

Finestra 3. Formalizzazione GOMS per il raggiungimento dello scopo: attiva l'icona.

L'esempio è tratto dall'applicazione di questa metodologia alla valutazione di OFFICE 241, un sistema informativo sviluppato da INSIEL S.p.A. per la gestione degli atti e la trasparenza dell'azione amministrativa (De Angeli, 1994).

L'analisi GOMS rappresenta la descrizione delle conoscenze che l'utente deve possedere per eseguire specifici compiti tramite sistemi informatici. Questa metodologia permette di evidenziare molteplici aspetti relativi alle dimensioni di efficacia e alcuni relativi all'efficienza (Kieras, 1988).

Consideriamo innanzitutto alcuni modi con cui il modello permette di valutare l'efficacia. La scomposizione del compito in scopi e sotto-scopi permette di evidenziare la *naturalità e logicità* delle procedure di soluzione previste dal sistema. La *completezza* del sistema viene verificata tramite il principio della necessità di un metodo per ogni scopo. La precisa formalizzazione di ogni possibile metodo permette di valutare la *linearità o pulizia*

della progettazione, come anche di verificare l'esistenza di chiare e ben determinate regole di selezione fra metodi diversi. L'assenza di tali regole implicherebbe l'inutilità di alcuni metodi. La *consistenza* dell'interfaccia può essere infine valutata mediante paragone fra metodi per il raggiungimento di scopi uguali in contesti diversi.

L'*efficienza* di un sistema viene inferita invece sulla base della brevità e del costo dei metodi, nonché di stime dei tempi necessari all'esecuzione di ogni operatore, inferite sulla base del Modello dell'Elaboratore umano. Bisogna ricordare, comunque, che tali stime sono adeguate soltanto in riferimento a utenti ideali, poiché anche i più esperti tendono a commettere errori o a utilizzare strategie di soluzione non ottimali. Non sono mancati, nel tempo, tentativi di estensione del modello in modo da includere comportamenti erronei e fenomeni dovuti all'apprendimento. I primi tentativi risalgono al modello CCT (*Cognitive Complexity Theory*), sviluppato da Polson e Kieras (1985), che produce regole *se-allora* a partire da un'analisi GOMS. Nel modello CCT il tempo di apprendimento viene inferito sulla base della lunghezza dei metodi, vale a dire del numero di operazioni che formano il metodo. La qualità della stima dipende dal livello degli operatori presenti nella formalizzazione. Le stime migliori si ottengono quando gli operatori appartengono al livello più basso, cioè alle *primitive standard*, per esempio l'azione di digitare una lettera. Il metodo permette di prevedere solamente il tempo necessario all'apprendimento di sequenze di azioni, dal momento che le modalità di esecuzione delle singole azioni vengono ipotizzate come note. Per una dettagliata bibliografia dei lavori successivi si veda l'interessante rassegna di Olson e Olson (1990).

I vantaggi della valutazione basata su analisi del compito sono legati alla sua oggettività e ai bassi costi di esecuzione. La previsione dei comportamenti erronei e dei fenomeni di apprendimento costituisce invece il suo limite. Di conseguenza, tale metodologia è più idonea alla valutazione di interfacce destinate a un uso continuativo, che di quelle destinate a un'utenza occasionale (Archer e Yuan, 1995). Essa permette infatti di misurare l'usabilità in termini di complessità del sistema, ma non di prevedere l'effettiva prestazione degli utenti (De Angeli, 1994). Un'ulteriore difficoltà risiede nella formalizzazione del comportamento di sistemi interattivi flessibili per il quale non sono disponibili strumenti adeguati (§ 1.5). In tutti questi casi, il ricorso alla valutazione con osservatori reali appare la scelta più idonea.

3.4. Valutazione con osservatori reali

Per eleganza stilistica, in questo lavoro, si è preferito tradurre l'espressione inglese *user-based evaluation* come valutazione con osservatori reali. Tale metodo costituisce l'approccio empirico alla misurazione dell'usabilità. In questo contesto, il termine *reale* va inteso in contrapposizione al concetto di utente *ideale* della valutazione mediante analisi del compito.

La valutazione basata sugli utenti implica costi maggiori rispetto alle altre forme di valutazione, ma solitamente offre risultati superiori, permettendo studi approfonditi del comportamento di interazione, come analisi dell'errore, delle strategie di recupero e della fase di apprendimento (Nielsen, 1993). Inoltre, essa consente di verificare la soddisfazione dell'utente che tanto influenza il successo commerciale di un sistema. Per quanto riguarda i sistemi interattivi flessibili poi, la valutazione empirica è attualmente l'unica metodologia affidabile.

La valutazione con osservatori reali può essere subordinata a due obiettivi principali, che richiedono approcci profondamente diversi:

- valutazione comparata
- valutazione singola

Il primo tipo di valutazione viene tipicamente utilizzato in fase di progettazione parallela (§ 3.1.3) e prevede il confronto dell'usabilità di più prototipi diversi. Il secondo è utilizzato invece durante la successiva fase di progettazione iterativa e serve esclusivamente a verificare l'usabilità di un particolare prototipo. La distinzione è fondamentale. Nel primo caso è necessario allestire un vero e proprio *setting* sperimentale, definendo i fattori di interesse ed eventualmente formulando le ipotesi sperimentali. Ciò richiede il rispetto delle tradizionali metodologie sperimentali psicologiche relative alla manipolazione fattoriale e al numero minimo di soggetti necessari per inferenze statistiche. Nel secondo caso invece è necessario esclusivamente evidenziare il massimo numero di problemi di usabilità. A questo fine cinque utenti possono essere sufficienti. Essi permettono infatti di evidenziare circa lo 80% dei difetti, mentre per migliorare tale risultato è necessario moltiplicare il numero del campione (Virzi, 1992; Lewis, 1994).

Le misure comportamentali possono essere raccolte tramite vari metodi. Uno dei più usati è il *logging*, la registrazione automatica di ogni azione prodotta dall'utente e dal computer durante l'interazione (van Nes e van Itegem, 1993; Nielsen, 1993). Il suo vantaggio è legato alla non intrusività dell'osservazione, lo svantaggio alla necessità di gestire grandi masse di dati. Tale problema si riferisce sia a difficoltà informatiche di

gestione dell'informazione, sia alla complessità della selezione degli eventi salienti in fase di analisi. Inoltre, il metodo permette solo di conoscere cosa è successo in un dato momento dell'interazione, non perché un dato comportamento sia stato eseguito. La mancanza di informazione sulle intenzioni dell'utente rende spesso difficile la comprensione degli errori.

Lo studio dei processi cognitivi sottostanti l'azione viene effettuato tramite la tecnica del *think-aloud* (Ericsson e Simon, 1984). Agli utenti è chiesto di verbalizzare ogni loro comportamento durante l'esecuzione del compito. L'interazione è registrata e in seguito sottoposta a analisi dei protocolli (Brouwer-Janse, 1991; Benysh, Koubek e Calvez, 1993). Secondo Nielsen (1994), il numero ideale di partecipanti per evidenziare con costi minimi e quantità soddisfacente di difetti di usabilità si aggira intorno a quattro o cinque. Il vantaggio fondamentale e specifico di tale tecnica è la raccolta di misure qualitative relative alle strategie di soluzione del compito. Gli svantaggi sono numerosi. Innanzitutto non è chiaro l'effetto della verbalizzazione su procedimento di soluzione: sono stati evidenziati dei casi in cui la verbalizzazione compromette dei processi cognitivi (Schooler e Engstler-Schooler, 1990; Schooler, Olson e Brook, 1993). Si tratta infatti di un metodo particolarmente intrusivo che può aumentare anche in modo pesante il carico cognitivo dell'utente e creare un contesto estremamente diverso da quello naturale. Inoltre, la sua esecuzione è piuttosto complessa, sia per il partecipante, che deve verbalizzare anche e specialmente nei momenti di difficoltà, sia per lo sperimentatore, che deve verificare l'effettiva verbalizzazione di ogni comportamento ed eventualmente stimolare l'utente. Per risolvere questi problemi il *think-aloud* può essere applicato come valutazione retrospettiva (Nielsen, 1993). L'utente utilizza il sistema normalmente e solo in seguito, di fronte alla completa videoregistrazione della sessione, egli è stimolato a verbalizzare le sue strategie. Il problema principale della variante è che spesso le persone non sono in grado di spiegare retrospettivamente le loro azioni. Un'ulteriore variante è denominata *codiscovery learning* (Kennedy, 1989) o *constructive interaction* (O'Malley, Draper e Riley, 1984). Entrambe le tecniche prevedono che due utenti partecipino contemporaneamente alla valutazione e si comunichino di volta in volta le loro intenzioni. Il vantaggio deriva da un *setting* più ecologico, lo svantaggio è invece legato all'aumento di possibili variabili confondenti legate alla presenza dell'altro partner umano.

Le misure soggettive possono essere raccolte tramite questionari o interviste strutturate. Il vantaggio dei questionari risiede nella possibilità di analisi statistiche più sofisticate. Numerosi strumenti psicometrici sono attualmente disponibili per la misurazione di aspetti diversi della soddisfazione dell'utente (Bailey e Pearson, 1983; Doll e Torkzadeh, 1988; Shneiderman, 1992²⁹), anche se si sente ancora la mancanza di

²⁹ Tale autore ha sviluppato il QUIS, Questionnaire for User Interface Satisfaction, di cui una versione italiana, attualmente in fase di taratura, è presentata in De Angeli, 1994.

rigorose verifiche di attendibilità, interna, esterna e di costrutto di tali strumenti. Uno degli ostacoli principali al rigore metodologico dipende dal fatto che tutte le dimensioni dell'usabilità sono *sistema specifiche*, vale a dire soggette a cambiamenti a seconda dell'oggetto valutato e del compito. Questo richiede continui adattamenti e revisioni degli strumenti, se non addirittura formulazioni *ex novo*, in contesti di ricerca il cui interesse principale non risiede nella validazione di questionari e che necessitano risultati in tempi rapidi.

3.4.1. Carenze metodologiche

L'aspetto metodologico della ricerca empirica è spesso sottovalutato. Non a caso, in uno dei pochi lavori specifici, tali aspetti vengono definiti *only means to ends* (Landauer, 1988; il problema della validità metodologica in generale è brevemente affrontato anche da Nielsen, 1993 e da Tague-Sutcliffe, 1992 relativamente a valutazioni di sistemi di *Information Retrieval*).

Il generale disinteresse verso la componente metodologica può essere attribuito a più fattori specifici della HCI, quali: (a) la natura dinamica, che richiede un costante aggiornamento delle metodologie (§ 2); (b) la natura applicativa, che richiede risultati veloci a prezzi contenuti (§ 2.2.1); (c) la disattenzione della psicologia accademica che dovrebbe rappresentare il settore più interessato a tali studi (§ 2.3.1); (d) la disattenzione per l'aspetto metodologico da parte dei progettisti, che spesso conducono di persona le valutazioni (Landauer, 1988).

Come abbiamo precedentemente visto, a livello metodologico gli sforzi si sono concentrati prevalentemente sulla definizione di criteri di minimo per ridurre il costo delle prove empiriche nelle valutazioni singole. Grande attenzione è stata riposta anche a un aspetto specifico della procedura sperimentale: la presentazione di scopi e obiettivi della ricerca. Attualmente si ritiene che tali presentazioni debbano essere elaborate in modo da ridurre eventuali *bias* dovuti alla sensazione di essere sotto giudizio o al risultato della prestazione personale. A tale scopo Nielsen (1993) propone un lungo elenco di direttive. I punti più interessanti sono: (a) necessità di presentare la valutazione come un lavoro finalizzato al miglioramento del prodotto informatico e non alla verifica di abilità personali; (b) necessità di sottolineare al massimo l'importanza del giudizio dell'utente, rendendolo così partecipe al ciclo di progettazione. Un'ulteriore strategia per incentivare la motivazione e diminuire eventuali fenomeni di dissonanza cognitiva consiste nel pagare i partecipanti (De Angeli, 1991).

Un aspetto metodologico estremamente importante, ma assolutamente trascurato, riguarda la scelta del disegno sperimentale (Nielsen, 1993). A causa della dispendiosità delle ricerche solitamente viene privilegiato il disegno *within-subject* (Archer e Yuan,

1995). Una veloce analisi delle valutazioni riportate da due volumi dell'*International Journal of Human-Computer Studies* del 1996 (volume 44 e 45) ha permesso di evidenziare che il rapporto fra tale disegno e quello *between-subjects* è di 8 a 3, e che solo 3 ricerche si basavano su disegni misti.

La manipolazione entro i soggetti ha il vantaggio di diminuire la variabilità individuale, ma pone il problema di una serie di variabili confondenti, quali l'effetto di apprendimento, stanchezza, (se tutte le prove vengono eseguite in successione), di storia (se le prove vengono eseguite in giornate diverse) o di regressione statistica, per cui quando un partecipante ha fornito una prestazione ai valori estremi in una prova, in quella successiva tenderà a ottenere un punteggio più vicino ai valori medi. Inoltre, se dopo ogni prova si richiede all'utente di fornire delle misure soggettive i limiti del disegno *within-subject* diventano ancora più evidenti, coinvolgendo anche l'effetto delle prove sul giudizio.

Una strategia suggerita per diminuire l'effetto di apprendimento consiste nel proporre compiti diversi nelle varie sessioni (Ingwersen, 1994). Il problema principale di tale metodo è la necessità di dimostrare a priori l'equivalenza dei compiti eseguendo sofisticati studi pilota. La difficoltà di soluzione spinge solitamente a ignorare il problema. Senza la conferma empirica, un disegno a compiti disomogenei deve essere considerato come un disegno a due fattori Compito * Sistema, cosa solitamente evitata. Inoltre, la definizione di compiti equivalenti appare particolarmente difficile nel caso di sistemi interattivi flessibili per l'esecuzione di compiti complessi. Un esempio pratico, tratto dalla mia personale esperienza nel settore, permette di comprendere la difficoltà nella scelta di adeguati disegni sperimentali. Un paio di anni fa mi sono trovata a collaborare alla valutazione di un sistema di *Information Retrieval* (IR), in cui dopo lunghe discussioni e studi pilota era stato scelto un disegno sperimentale a misure ripetute con compiti diversi, procedura ampiamente consigliata dalla letteratura del dominio. I partecipanti dovevano interagire con due versioni di IR eseguendo ogni volta un compito diverso. Malgrado l'attenzione riposta alla verifica dell'equivalenza, in fase di analisi statistica è emerso che un compito era significativamente più facile dell'altro e, cosa ancora più problematica, un'interazione significativa Compito * Sistema. Poiché la valutazione, basata prevalentemente sull'analisi dell'apprendimento, non prevedeva il bilanciamento del fattore Sistema, il vantaggio derivato dal reclutamento di un numero inferiore di partecipanti è stato annullato dallo svantaggio della difficoltà di inferenza statistica.

Ovviamente la scelta di un adeguato disegno sperimentale dipende dal particolare contesto di analisi e dalle finalità della ricerca. Personalmente comunque sono portata a preferire il disegno *between-subject*, in quanto permette di raccogliere il massimo numero di variabili per ogni partecipante evitando di stancarlo eccessivamente. A causa delle numerose variabili in gioco nell'interazione flessibile, tale disegno richiede però lo

sviluppo di adeguati criteri di selezione dei partecipanti, sulla cui base definire il campione che deve poi essere assegnato alle diverse condizioni in modo casuale (Nielsen, 1993; § 6.2.2). Riprenderemo questa problematica più volte nel corso dell'esposizione dei risultati delle due simulazioni presentate nella parte sperimentale, per adesso occupiamoci di descrivere la tecnica di valutazione più comunemente usata per l'analisi dei sistemi interattivi flessibili: la simulazione.

3.5. *La simulazione come strumento conoscitivo*

Nella sua accezione comune il termine simulazione significa manifestare sentimenti, intenzioni o condizioni non corrispondenti a realtà per secondi fini più o meno riprovevoli (Devoto e Oli, 1990). In ambito scientifico il termine perde ogni connotato negativo per tornare al significato originario del latino *simulare*, letteralmente rendere simili. La scienza in generale è ricorsa spesso alla simulazione, intesa come riproduzione semplificata della realtà. In psicologia, per esempio, ogni esperimento *simula* il fenomeno oggetto di studio, nel senso che lo riproduce artificialmente. Il fenomeno viene simulato per essere analizzato al riparo dall'influenza di altre variabili che nella realtà tendono a manifestarsi in concomitanza a esso. In informatica un *simulatore* è un programma che imita le caratteristiche di un sistema di elaborazione e le riproduce all'interno di un sistema diverso. Accanto a questa accezione, attualmente si sta affermando anche il concetto di *simulazione* al computer, riproduzione di un fenomeno o di un comportamento reale tramite l'elaboratore.

Nella HCI il termine *simulazione* si riferisce a una tecnica di validazione di prototipi (*rapid prototyping*; Nielsen, 1993), che può riguardare la riproduzione tramite il computer di particolari strumenti interattivi o la riproduzione tramite operatore umano di particolari funzionalità di elaborazione. Nel primo caso per esempio si valuta una tastiera riproducendola sullo schermo di un computer (Archer e Yuan, 1995). In queste situazioni la simulazione ha dimostrato di fornire risultati paragonabili alla valutazione dello hardware fisico, in modo molto più economico e flessibile.

Nel secondo caso, invece, la simulazione è più complessa e rappresenta l'unica soluzione possibile di un paradosso: è impossibile progettare un sistema di interazione utente-computer senza conoscere le caratteristiche peculiari di tale interazione, ma è altrettanto impossibile studiare l'interazione senza disporre del sistema informatico capace di instaurarla.

L'importanza della simulazione come mezzo per la definizione di un supporto teorico alla progettazione è ampiamente testimoniata dalla storia dello sviluppo dei sistemi di comprensione del linguaggio naturale (Guyomard e Siroux, 1988; Jönsson e Dählback, 1988; Fraser e Gilbert, 1991). La maggior parte dei primi prototipi furono

costruiti unicamente sulla base dei pregiudizi teorici e delle intuizioni dei loro progettisti. Per questo motivo alla resa dei conti spesso si sono rivelati estremamente complessi e troppo limitati, in altre parole praticamente inusabili.

Attualmente sono disponibili vari metodi oltre alla simulazione finalizzati all'identificazione e alla previsione del *coverage* linguistico di un sistema, vale a dire del set di input che dovrebbe essere in grado di maneggiare durante l'interazione con l'umano. Analizziamo pregi e difetti dei metodi principali.

- Il metodo *test suites* prevede la generazione su base teorica di un campione rappresentativo di tutti i fenomeni linguistici che il sistema, una volta operante, dovrebbe essere in grado di gestire. La difficoltà principale è legata alla limitatezza dei fenomeni evidenziabili (Whittaker e Stenton, 1989). Infatti, mentre le teorie sintattiche forniscono elaborate tassonomie, le teorie pragmatiche non hanno ancora raggiunto un adeguato livello di formalizzazione.
- Il metodo *carta e matita* si basa sulla raccolta empirica di un campione rappresentativo di espressioni linguistiche, chiedendo a delle persone di scrivere la frase necessaria al raggiungimento di un dato scopo. Il problema principale riguarda la rappresentatività dei compiti che devono essere risolti da un'unica richiesta. Anche questo approccio è quindi destinato a sottovalutare tutti i fenomeni pragmatici e interattivi.
- L'*analisi dell'interazione utente-utente* si svolge studiando la comunicazione prodotta da due persone per risolvere un dato compito. L'inadeguatezza del metodo è legata alla differenza strutturale fra questo tipo di interazione e quella supportata dalla macchina (§ 2.1.1).
- L'*analisi sul campo* implica la valutazione di un prototipo effettivamente esistente. Il problema principale riguarda la fragilità e la difficoltà d'uso delle applicazioni disponibili che possono venire usate solo da un'utenza addestrata e consapevole dei limiti. Il comportamento linguistico risulta dunque inevitabilmente viziato dalle capacità del sistema.

3.5.1. Simulazione Wizard of Oz

La tecnica di simulazione *Wizard of Oz* (da ora in poi, WOZ), nota anche con l'acronimo PNAMBIC (*Pay No Attention to the Man Behind the Curtain*), è stata elaborata per analizzare l'interazione senza dovere realizzare completamente il sistema. La procedura è semplice. Come nel celebre romanzo *The Wizard of Oz* scritto da Baum nel 1900, un essere umano (il mago) gioca il ruolo del computer in una simulazione di conversazione utente-computer. Ai partecipanti viene fatto credere di interagire con una macchina, mentre in realtà dietro all'interfaccia del computer c'è il mago. Egli tramite un

computer collegato in rete con quello del partecipante, gestisce l'interazione decifrando gli input e generando gli output. Solitamente la simulazione viene eseguita in laboratori di usabilità, composti da due stanze separate. Quella di osservazione, in cui sono situati vari strumenti di monitoraggio è dedicata al mago, o meglio ai maghi. Infatti è consigliabile suddividere tale ruolo fra due persone: una addetta alla gestione dell'interfaccia, l'altra di supporto per l'interpretazione degli input e la generazione degli output, nonché per il monitoraggio degli strumenti di controllo. Il costo economico di tale procedura è ampiamente ricompensato dall'aumento della validità della simulazione (De Angeli, 1991; De Angeli *et al.*, 1996).

La tecnica WOZ, pur non essendo una novità³⁰, è stata utilizzata in un numero limitato di lavori. E' infatti una metodologia difficile e dispendiosa. Secondo Fraser e Gilbert (1990) la qualità della simulazione è funzione di tre requisiti fondamentali:

- *realizzabilità*: adeguatezza del mago a simulare il comportamento del sistema
- *formalizzazione del modello dialogico del sistema*: specificazione delle capacità interattive
- *credibilità*: per tutta la durata dell'interazione il soggetto non deve scoprire il trucco

La simulazione di sistemi multimodali aumenta il numero di problematiche (Salber e Coutaz, 1995), sia per la minore esperienza pratica, sia per la necessità di simulare sistemi più complessi. Questo richiede alcuni cambiamenti nella tecnica tradizionale che coinvolgono il ruolo del mago, il software e lo hardware di supporto alla simulazione. Dal punto di vista del mago il compito diventa più impegnativo e quindi è indispensabile coinvolgere più osservatori³¹. Dal punto di vista tecnologico invece è necessario l'uso di protocolli di comunicazione elettronica flessibili, capaci di gestire più modalità di interazione e postazioni di lavoro. Particolarmente impegnativa diventa anche la gestione del file di *log*. I dati provenienti da più modalità devono essere sincronizzati per permettere di ricostruire lo svolgimento temporale dell'interazione.

Il problema fondamentale sia per il mago che per la componente tecnologica è ridurre al minimo i tempi di risposta e di eventuali errori. È stato dimostrato infatti che intervalli lunghi tendono a influenzare il comportamento interattivo e il giudizio dell'utente (Fraser e Gilbert, 1991). Per superare questi problemi Oviatt e colleghi (1992) propongono la simulazione *semi-automatica*. In parole semplici la loro idea invita a

³⁰ I primi esperimenti risalgono ai lavori di Malhorta del 1975 e di Thomas 1976; (rif. bib. in Jöhnsson e Dählback, 1988).

³¹ Questo può essere vero anche per la simulazione di sistemi unimodali. La differenza fondamentale consiste nella divisione dei lavori. Nel caso di sistemi multimodali complessi ogni mago può gestire da una postazione di lavoro individuale una specifica modalità comunicativa. Il problema diventa allora la sincronicità e la collaborazione fra i maghi.

dedicare la massima attenzione alla scelta dei compiti e alle interfacce dei sistemi utilizzati da mago e utente. Compiti chiaramente strutturati permettono di prevedere il comportamento dell'utente e di limitarne la variabilità. Lo stesso effetto è prodotto da un'interfaccia utente strutturata. L'interfaccia del mago, invece, deve venire dotata di filtri automatici che limitino il verificarsi di azioni errate e di una serie di frasi standard che coprano il maggior numero di possibili di output del sistema.

Torneremo sui problemi della simulazione nella parte sperimentale per adesso occupiamoci della descrizione dell'approccio globale.

4. L'approccio globale alla HCI

L'approccio globale si propone come un contributo metodologico alla valutazione con osservatori reali (§ 3.4). Il suo dominio di applicazione è l'analisi di sistemi interattivi altamente flessibili, vale a dire del processo dialogico prodotto dall'interazione fra utenti e artefatti socio-cognitivi. Il suo utilizzo è finalizzato a valutazioni di usabilità di prototipi (obiettivo applicativo a breve termine) e allo sviluppo di una psicologia della HCI, capace di comprendere e spiegare l'oggetto di studio nella sua complessità (obiettivo teorico a lungo termine). L'obiettivo applicativo dunque costituisce la base di quello teorico, in quanto fornisce dati e osservazioni. L'obiettivo teorico poi si articola su due aspetti complementari. Il primo coincide con l'elaborazione di un apparato metodologico sulla cui base sviluppare modelli *interpretativi* e *descrittivi* del comportamento dei sistemi interattivi. Il secondo corrisponde alla traduzione dell'apparato teorico così prodotto in direttive e metodologie direttamente utilizzabili dai progettisti.

Alla base dell'approccio globale si situano le riflessioni sulla natura della HCI e sulle intrinseche difficoltà del contributo psicologico in questo dominio, espone nei capitoli precedenti. Per chiarezza esplicativa, riassumiamo brevemente i concetti fondamentali.

- Il dialogo flessibile è mediato da artefatti socio-cognitivi: nell'interazione vengono coinvolte sia capacità cognitive che abilità sociali (§ 1.4).
- L'allargamento del campo dialogico complica la valutazione (1.5).
- La HCI è una disciplina dinamica: il suo apparato teorico e metodologico deve continuamente essere adattato a un oggetto di studio *in fieri* (§ 2).
- Il trapianto diretto di teorie e metodi psicologici è inadeguato: è necessario sviluppare una specifica psicologia della HCI (§ 2.3.1).
- Secondo la *prospettiva interazionale* la psicologia della HCI richiede il coinvolgimento di più settori psicologici (§ 2.4.2).
- L'usabilità richiede conoscenza specifica dell'interazione: la qualità dipende da caratteristiche del computer, dell'utente, del compito e del contesto (§ 3.1).

4.1. *La nuova proposta*

Alla base dell'approccio globale si situa la definizione di interazione come *totalità dinamica e complessa prodotta dall'azione sinergica di quattro entità: utente, computer, compito e ambiente*. È chiaro che tale idea deve molto alle proposte precedenti (§ 2.2.1). Il suo valore innovativo consiste piuttosto nel fornire un'organizzazione alle specifiche entità, riadattando al particolare dominio di analisi alcuni concetti paradigmatici della psicologia della gestalt. Definire l'*interazione* come *totalità* implica concepirla come qualcosa di diverso e superiore alla somma delle singole entità che la determinano. Si noti che sotto questo aspetto ci si contrappone nettamente alla proposta di Booth (1989) che definiva la HCI come somma di specifiche aree di ricerca.

Da un punto di vista metodologico, la nuova definizione induce a sostenere che la comprensione dell'interazione non può derivare da analisi separate delle singole entità nè di rapporti, parziali, quali, per esempio, lo scambio fra utente e computer. Secondo la teoria gestaltista infatti ogni entità è caratterizzata in modo diverso se considerata singolarmente o inserita nel tutto; inoltre una stessa entità inserita in due diverse totalità può assumere caratteristiche diverse. Il primo postulato spiega l'insuccesso del puro trapianto dell'apparato teorico cognitivista, il secondo invece di quello della psicologia sociale (§ 2.3.1). Secondo la nuova prospettiva, l'interazione deve essere spiegata tramite teorie di campo, che descrivano i principi di unificazione formale con cui le singole parti danno luogo allo specifico risultato.

In questo senso l'approccio globale si allontana decisamente dalla tradizionale metodologia ergonomica. La distanza deriva dal rifiuto del *principio di modularità*, su cui l'approccio cognitivista basa l'analisi di sistemi complessi (Simon, 1969). È questo un passaggio obbligato, indotto dallo sviluppo tecnologico. Infatti, in contesti di interazione altamente vincolata l'approccio modulare può fornire risultati soddisfacenti se applicato all'analisi del comportamento dell'essere umano. Ciò è dovuto alla minore complessità della situazione: le variabili che determinano il comportamento sono quantitativamente minori e qualitativamente più semplici (Landauer, 1987). Eppure, lo stesso approccio risulta inadeguato se applicato al computer. L'usabilità di una singola funzione valutata isolatamente non è generalizzabile se la funzione viene affiancata da altre (Landauer e Galotti, 1984; Landauer, 1988). Infatti, si possono creare situazioni di interferenza che aumentano il carico cognitivo richiesto all'utente, per cui anche la più semplice procedura a volte risulta estremamente difficoltosa. A maggior ragione, l'approccio modulare appare inadeguato allo studio di sistemi interattivi flessibili. L'aumento del numero e della complessità delle variabili in gioco, nonché la loro natura intrinsecamente correlata rendono il processo di scomposizione problematico e inattendibile.

È importante notare fin da ora, che l'approccio globale non esclude la possibilità di eseguire studi finalizzati alla comprensione di specifiche sottoparti del processo. Piuttosto esso richiede di controllare sempre ogni possibile sorgente di variabilità: se l'interesse si orienta verso lo studio del rapporto fra due specifiche entità occorre mantenere costanti le altre.

4.1.1. Pragmatica del metodo

L'esecuzione di una valutazione affidabile è subordinata a un lungo lavoro di preparazione in cui sono coinvolte competenze diverse. Da un punto di vista operativo, il metodo di valutazione proposto si sviluppa su sette fasi.

1. *Definizione degli obiettivi della ricerca.* Innanzitutto è necessario stabilire esattamente cosa si vuole studiare e perché. Ciò implica la scelta del tipo di valutazione, (comparata o singola), del metodo (valutazione di sistemi completi, prototipi reali o simulazione) nonché un accurato esame di fattibilità della ricerca e della letteratura relativa al fenomeno in esame. In questa fase vengono anche formulate a livello generale eventuali ipotesi di lavoro.
2. *Caratterizzazione delle quattro entità.* Per ognuna di esse è necessario identificare i principali fattori che possono influenzare l'interazione (da ora in poi, *determinanti* dell'interazione) e stabilirne le correlazioni. Ogni determinante deve essere poi operazionalizzata, definendo la natura e i criteri di misurazione.
3. *Selezione degli attributi di usabilità.* A seconda del tipo di valutazione e delle determinanti evidenziate è necessario definire esattamente le variabili dipendenti (attributi di usabilità) sulla base di un accreditato modello (§ 3.1), i criteri di minimo (§ 3.1.3) e selezionare i metodi di misurazione (§ 3.4). Dopodiché vanno formulate nel dettaglio le ipotesi di lavoro.
4. *Definizione dello scenario.* A questo livello si elabora il metodo sperimentale. Ciò richiede: la selezione di un campione rappresentativo nonché l'elaborazione delle procedure e dei compiti di valutazione. La scelta del campione deriva da accurati compromessi fra obiettivi della ricerca, caratteristiche del sistema e risorse disponibili. Per ogni compito invece occorre definire accurati criteri di correttezza, produrre dettagliate formalizzazioni e valutarne la valenza ecologica sulla base del sistema oggetto di valutazione.

5. *Breve studio pilota.* A differenza della tradizionale metodologia psicologica, dove lo studio pilota serve prevalentemente a raffinare le ipotesi sperimentali, nelle valutazioni di usabilità il suo scopo è totalmente subordinato alla verifica dello scenario.
6. *Esecuzione della valutazione.* Lo studio deve essere condotto con grande attenzione e sensibilità in modo da evitare di inserire ulteriori variabili di disturbo nella già complessa situazione.
7. *Analisi dei risultati.* Tipicamente le valutazioni di usabilità forniscono una mole impressionante di dati che spaziano da misure quantitative a osservazioni qualitative, da indici oggettivi a giudizi soggettivi. In fase di analisi è dunque importante considerare tutta l'informazione disponibile, sia per fornire un quadro completo della situazione sia, per evidenziare eventuali incongruenze che possono riflettere errori procedurali. È doveroso inoltre verificare sempre l'effettivo successo di eventuali manipolazioni sperimentali.

Nei prossimi capitoli verrà fornita una descrizione dettagliata delle fasi 2 e 7 che corrispondono ai punti in cui l'approccio globale, per il suo peculiare interesse all'interazione nella sua totalità, si differenzia maggiormente dalla tradizionale procedura ergonomica.

4.2. *Caratterizzazione delle entità*

Il numero e la qualità delle determinanti evidenziate nella fase di caratterizzazione dipende dal livello desiderato di attendibilità e generalizzabilità dei risultati, nonché dalle particolari caratteristiche delle entità coinvolte nella valutazione. Abbiamo già visto infatti, che il numero di determinanti riferibili all'utente è funzione del tipo di sistema informatico testato: gli artefatti cognitivi tendono a coinvolgere quasi esclusivamente capacità mentali, gli artefatti socio-cognitivi anche caratteristiche individuali e abilità sociali.

La Figura 4 rappresenta schematicamente le determinanti principali per ognuna delle quattro entità.

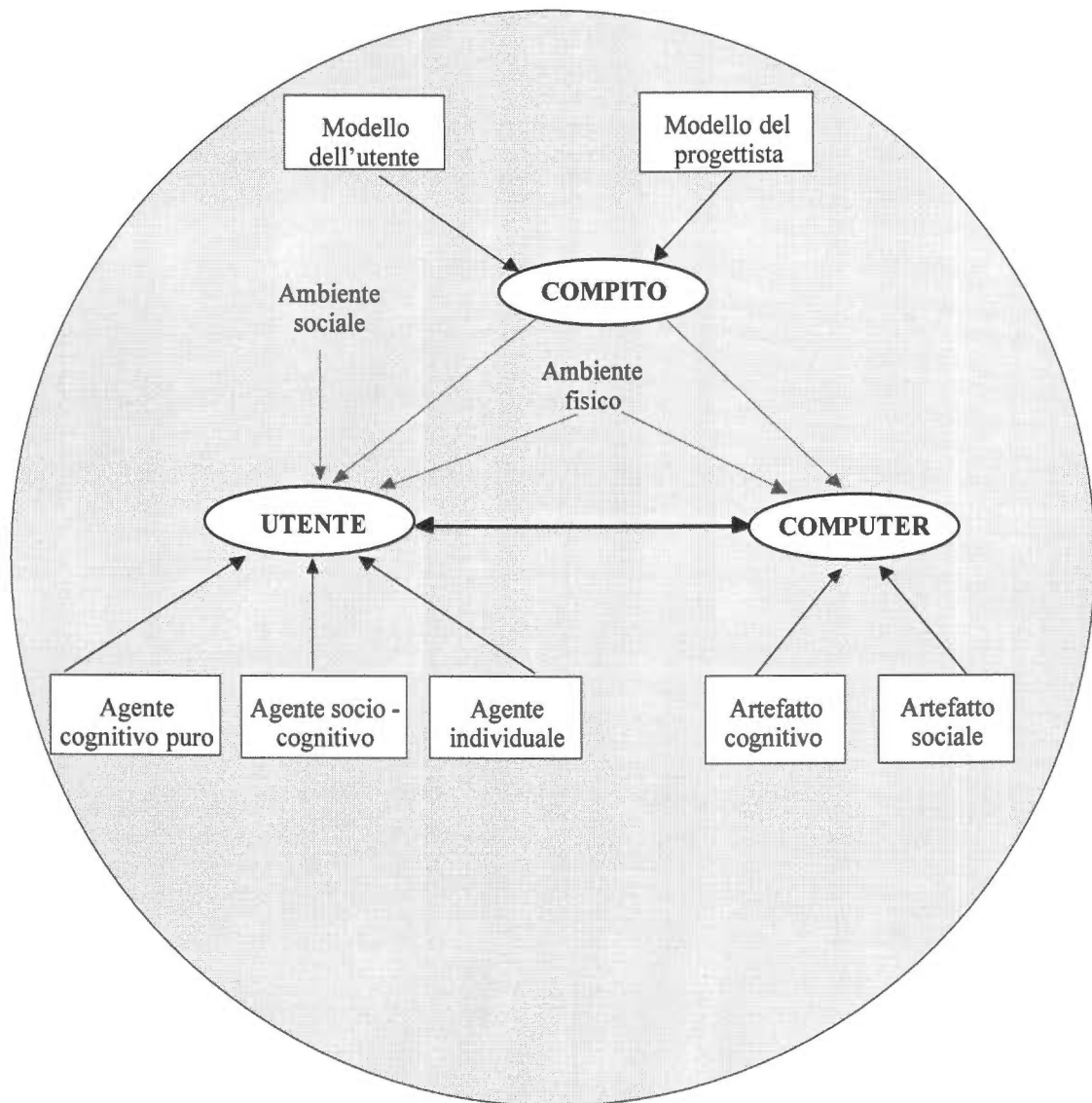


Figura 4. Determinanti dell'interazione secondo l'approccio globale.

L'interazione ha luogo in un determinato ambiente (graficamente rappresentato dal cerchio grigio) al fine di eseguire uno specifico compito. Utente e computer sono definiti entità dirette, nel senso che agiscono direttamente sull'interazione. Ambiente e compito sono invece entità indirette, il loro effetto (graficamente rappresentato dalle linee più chiare) è mediato dalle entità dirette.

4.2.1. Ambiente

L'ambiente è l'entità cui storicamente è stata dedicata minore attenzione anche se, negli ultimi anni, si evidenzia un lieve cambiamento di tendenza (Lewis, 1990). Eppure, il contesto in cui la valutazione viene effettuata può pesantemente influenzare la generalizzabilità dei risultati. La soluzione ideale è dunque il laboratorio di usabilità, un ambiente di osservazione del comportamento che riproduce il luogo in cui i sistemi informatici vengono usati, tipicamente dunque un ufficio. L'importanza dell'ambiente aumenta quando si vogliono misurare variabili soggettive, quali soddisfazione dell'utente, frustrazione e stress indotto dall'interazione. Infatti, il fatto stesso di essere direttamente

osservati durante lo svolgimento di un compito complesso e poco familiare può indurre stress e tensione e quindi inquinare i risultati della prova. Per questo tipicamente un laboratorio di usabilità è composto da due stanze messe in collegamento da uno specchio unidirezionale, attraverso cui viene osservata la situazione interattiva.

Secondo l'approccio globale le determinanti interattive dell'ambiente sono il *contesto fisico* e il *contesto sociale*. La prima agisce direttamente sia sull'utente che sul sistema e attraverso questi tende a influenzare il compito. Si consideri, per esempio, un ambiente rumoroso. Tale contesto può compromettere anche seriamente le prestazioni di un sistema di riconoscimento del parlato o di un utente che affronta un problema impegnativo. In entrambi i casi il compito tende a diventare più difficile.

Il contesto sociale invece agisce direttamente solo sull'utente e tramite lui influenza le altre entità. Secondo la teoria della *categorizzazione del sé* (Turner, Hogg, Oakes, Reicher e Wetherell, 1987), le persone non sono caratterizzate da un sé individuale fisso, ma da una varietà di categorie del sé, ciascuna delle quali può divenire saliente a seconda del contesto sociale. In questa visione, il concetto di sé è dunque definito come una struttura cognitiva formata da varie componenti che corrispondono alle diverse identità che la persona possiede. Tali componenti si articolano su un continuum di identità e possono funzionare in modo relativamente indipendente influenzando il comportamento. Ai fini dell'approccio globale sono particolarmente importanti i due poli estremi del continuum: (a) *l'identità sociale*, vale a dire il concetto di sé come membro di un gruppo con cui l'individuo si identifica; (b) *l'identità individuale*, il concetto di sé in quanto individuo unico e irripetibile. Quando la componente sociale è saliente tendiamo a comportarci secondo le norme sancite dal gruppo; altrimenti ci comportiamo più facilmente sulla base della nostra personalità. La salienza dell'identità sociale è considerata una variabile fondamentale per predire i processi socio-cognitivi che guidano il comportamento nella comunicazione mediata dal computer (Spears, Lea e Lee, 1990; Lea e Spears, 1991, 1992; Mantovani, 1993; Spears e Lea, 1992, 1994). E' lecito anche ipotizzare un suo effetto principale sull'utilizzo di sistemi per il lavoro di gruppo, sulla scelta dei siti da visitare navigando in Internet o di uno specifico vocabolario nell'interazione basata sul linguaggio naturale. Allo stesso modo il contesto sociale può influenzare la validità dei giudizi soggettivi. Questo fenomeno rappresenta una complicazione del tipico problema psicologico del *buon soggetto*. Se l'utente ritiene che lo sperimentatore sia in qualche modo coinvolto nella progettazione del sistema il suo giudizio è generalmente più positivo (Nielsen, 1993).

4.2.2. Compito

Nella visione dell'approccio globale il *compito* viene descritto come un esercizio di *problem solving* (Newell e Simon, 1972; Simon, 1975; Rumiati, 1984). Lo stato meta coincide con lo scopo per cui l'utente utilizza il computer. La soluzione dipende dall'esecuzione di particolari operazioni associate alla specifica conoscenza procedurale e descrittiva che l'utente possiede riguardo al problema e al modo di interagire con il computer.

Il lavoro sottostante la soluzione è suddiviso fra le due entità (utente e computer), lungo un continuum di allocazione di funzioni, che si sviluppa da un minimo a un massimo di automazione (Wærn, 1989). Tipicamente, la parte di pertinenza del sistema coinvolge la componente di elaborazione dell'informazione; quella di pertinenza dell'utente la definizione dello stato iniziale e la valutazione dello stato finale. La posizione sul continuum dipende dal tipo di compito (Wærn, 1989). I computer sono particolarmente adatti a eseguire procedure ripetitive e di precisione, risolvere problemi sulla base di algoritmi ed effettuare controlli di routine. Gli umani invece preferiscono compiti creativi e di bassa precisione, nonché problemi la cui soluzione dipende da intuizione e il monitoraggio attivo.

Secondo l'approccio globale le principali, determinanti interattive riferibili al compito sono il *modello dell'utente* e il *modello del progettista*. Il primo rappresenta la costruzione teorica sulla cui base l'utente agisce; il secondo quella implementata nel sistema, sulla cui base l'utente dovrebbe agire. La compatibilità fra tali modelli influenza direttamente la possibilità di successo della soluzione. Il modello dell'utente a sua volta è costituito da due determinanti: il *modello del compito prototipo* (insieme di strategie e conoscenze normalmente adottate per raggiungere la soluzione in condizioni naturali), il *modello del compito informatico* (insieme di strategie e conoscenze normalmente adottate per la soluzione del compito tramite computer). I due aspetti possono coincidere o essere molto diversi fra loro, ma entrambi tendono a influenzare il comportamento dell'utente.

4.2.3. Computer

Il computer è lo strumento tramite cui il compito viene risolto. L'interazione è dunque determinata dalla sua effettiva capacità di raggiungere gli scopi per i quali è stato progettato e dalla facilità di utilizzo. Nel primo senso l'elaboratore è definito come *artefatto cognitivo*, nel secondo come *artefatto sociale*.

L'artefatto cognitivo è costituito dalla parte esecutiva dello strumento. La definizione di tale determinante implica dunque la descrizione dello *hardware* (per esempio, risoluzione del display, velocità di elaborazione dell'informazione, memoria disponibile,

precisione degli strumenti di puntamento) e del *software* (algoritmi di soluzione, disponibilità dell'informazione). Quest'ultima sotto alcuni aspetti si sovrappone a una delle determinanti relative al compito, vale a dire al *modello del progettista*. L'artefatto sociale è invece costituito dall'*interfaccia*. La sua definizione prevede una dettagliata descrizione delle capacità dialogiche del computer.

La fase di caratterizzazione del computer assume particolare importanza nel caso di valutazioni basate su simulazione semi-automatica. In essa infatti deve essere dettagliatamente specificato e in qualche modo inventato il particolare sistema che si vuole riprodurre (§ 3.5.1). Questo compito prevede un'elaborata procedura. Innanzitutto bisogna stabilire il tipo di sistema da riprodurre a livello di artefatto cognitivo: una simulazione affidabile riproduce fedelmente lo hardware e le capacità del software previsti per il sistema finale. La definizione dell'artefatto sociale prevede invece l'elaborazione del modello del comportamento interattivo del sistema, vale a dire della sua abilità linguistica e comunicativa. Una prima formalizzazione viene elaborata combinando i risultati dei metodi alternativi alla simulazione, descritti nel capitolo precedente (§ 3.5).

Il metodo *test suites* è particolarmente efficiente se il campionamento viene sviluppato sulla scia di precedenti esempi di interazioni utente-computer. In questo modo si evita di concentrare l'attenzione su fenomeni poco probabili. Il metodo *carta e matita*, invece, risulta generalmente poco utile a causa della sua inadeguatezza a modellare situazioni interattive. Risultati migliori si possono ottenere da una sua variante (sviluppata e testata durante la progettazione delle simulazioni presentate in questo lavoro), in cui il metodo viene combinato all'analisi della *comunicazione utente-utente*. I due partecipanti interagiscono in una sessione di comunicazione mediata dal computer sincrona, o tramite un sistema di microfoni, a seconda del tipo di linguaggio che si vuole raccogliere. Si riproduce così una situazione interattiva da cui vengono filtrati buona parte dei fenomeni non verbali, estremamente informativi nello scambio faccia a faccia, ma estranei all'interazione utente-computer. Se poi uno dei partecipanti, d'accordo con lo sperimentatore, finge di essere straniero, si riproduce un contesto di comprensione non ottimale, più simile a quello da modellare. Si noti che le persone coinvolte in questa fase non dovrebbero poi essere chiamate a partecipare alla simulazione, in quanto potenzialmente poco ingenui. La formalizzazione così ottenuta viene poi raffinata nella fase di progettazione iterativa necessaria allo sviluppo dell'ambiente di simulazione.

4.3. Utente

La caratterizzazione dell'utenza viene considerata da più autori come un passo fondamentale per la comprensione dell'interazione (Nielsen, 1993; Coniglio e Paci, 1994; Tague-Sutcliffe, 1992). Eppure tuttora si evidenzia la mancanza di un apparato teorico capace di spiegare come le caratteristiche individuali agiscono sul processo interattivo. In generale gli sforzi si sono concentrati sullo studio dell'influenza di dati socio-anagrafici, quali sesso, titolo di studio (Sensales e Greenfield, 1991), status socio-economico (Parasumaran e Igarbia, 1990) e professione (Frantzve, Lips, Castor, Wilson, Kraus e Foster, 1992; Matheson, 1992). Tali variabili sono considerate indizi affidabili per l'identificazione di fasce a rischio di *emarginazione elettronica* e per la progettazione di sistemi a misura di utente (§ 1.1). Tuttavia il loro valore predittivo è oggetto di discussione. Come dimostreremo più avanti (§ 5), ciò deriva dal fondamentale errore della ricerca attuale, che consiste nell'aver analizzato il ruolo di tali variabili isolatamente da altre caratteristiche individuali, quali l'esperienza (Parasumaran e Igarbia, 1990; Marquié, Thon e Baracat, 1994).

In generale, si ritiene che gli utenti più anziani tendano a essere meno propensi dei giovani ad accettare e utilizzare nuove tecnologie (Parasumaran e Igarbia, 1990; Marquié *et al.*, 1994). A parità di accettazione, poi, la prestazione degli anziani è maggiormente influenzata da caratteristiche fisiche dell'interfaccia, quali la risoluzione grafica. Si ritiene, inoltre, che le donne abbiano un atteggiamento più negativo degli uomini nei confronti della tecnologia informatica e di conseguenza ne facciano un minore uso (Hess e Miura, 1985; Dambrot Watkins-Malek, Silling, Marshall e Garver, 1985; Popovich, Hyde e Zakrajsek, 1987). Secondo la *National Science Foundation* fra i laureati in informatica il rapporto uomini-donne è di 3 a 1 (Kantrowitz, 1994) e il gap sembra destinato a crescere. I corsi specialistici sono frequentati maggiormente da studenti maschi (Hess e Miura, 1985; Dambrot *et al.* 1985; Popovich *et al.*, 1987) che tendono a raggiungere prestazioni migliori (Fetler, 1985) e passano più tempo lavorando col computer rispetto alle loro coetanee (Popovich *et al.*, 1987). Il minore coinvolgimento delle donne in campo informatico ha trovato spiegazioni diverse che, a seconda del prevalente orientamento teorico, possono essere riassunte in due punti di vista principali: ipotesi *disposizionalista* e *situazionalista* (De Angeli, 1995).

L'ipotesi più popolare a prevalente orientamento disposizionalista riconduce le differenze nell'accettazione del computer a differenze nelle predisposizioni verso abilità logico-matematiche (Hess e Miura, 1985, Dambrot *et al.* 1985, Lockheed, 1985; Bandalos e Benson, 1990; Hawkins, 1985; Gilroy e Desai, 1986). Partendo dal presupposto che il computer venga prevalentemente percepito come strumento per la

manipolazione formale di simboli, il suo evitamento è interpretato come la generalizzazione di sentimenti negativi dal target originario (la matematica) al nuovo target (il computer). Varie ricerche hanno, infatti, dimostrato come la differenza fra i sessi sia particolarmente accentuata nella programmazione e tenda invece a scomparire in applicazioni quali la videoscrittura o la manipolazione di dati (Lockheed, 1985; Hess e Miura, 1985). Per quanto elegante, tale spiegazione presenta alcuni punti deboli che la rendono insufficiente a spiegare il fenomeno nella sua complessità. Innanzitutto l'esistenza di differenze sessuali nell'atteggiamento verso il computer è controversa. (Lockheed, 1985; Parasumaran e Igbaria, 1990; Coover e Delcourt, 1992; Frantzve *et al.*, 1992). Come dimostreremo empiricamente (§ 5), in campioni caratterizzati da un livello di esperienza omogeneo, le differenze fra uomini e donne tendono a scomparire. Inoltre, lo sviluppo delle interfacce grafiche avrebbe già da tempo dovuto ridurre la forza dell'associazione con la matematica e quindi porre fine al gap fra uomini e donne.

L'orientamento situazionalista, invece, attribuisce la minore presenza femminile nel mondo informatico ad alcune caratteristiche peculiari della società e in particolare alla distinzione fra ruoli sessuali. Numerose ricerche hanno dimostrato come lo stereotipo del computer costantemente trasmesso dalle pellicole cinematografiche, dalla TV e dalla stampa sia prevalentemente maschile (Ware e Stuck, 1985). L'incoraggiamento sociale e materiale dato a un maschio che si avvicina al mondo del computer sembra superiore a quello dato a una femmina (Hess e Miura, 1985). Inoltre, basta scorrere la lista dei titoli dei video giochi disponibili sul mercato (considerati un elemento critico nello strutturarsi delle motivazioni ad interagire con il computer) per rendersi conto dello squilibrio a favore degli interessi maschili (Hess e Miura, 1985; Ware e Stuck, 1985) che si riflette in un effettivo maggiore utilizzo da parte dei maschi (Lockheed, 1985).

L'approccio globale in generale accetta l'ipotesi situazionalista e di conseguenza ritiene che le variabili anagrafiche di per sè non siano sufficienti a caratterizzare l'utente (§ 5).

4.3.1. Un nuovo modo di concepire l'utente

Nella visione dell'approccio globale, l'utente viene definito come un soggetto in situazione che agisce sullo stato di cose percepito in vista di uno stato di cose concepito. Dunque, l'interazione è considerata funzione sia di caratteristiche oggettive di sistema, compito e ambiente, sia del modo in cui l'utente si rappresenta tali entità. In quest'ottica la risposta all'output di un sistema non viene considerata un puro atto cognitivo, ma il risultato del significato che l'utente attribuisce al dato output. Quanto più un sistema viene concepito in senso antropomorfo, tanto più l'attribuzione sarà mediata da schemi, processi inferenziali e stereotipi sociali.

Per definire le determinanti dell'interazione relative all'utente è utile partire dal modello dell'*attività umana* proposto da Norman (1983; Norman e Draper, 1986). Secondo tale autore, il comportamento di interazione si articola in sette stadi fondamentali (vedi Figura 5). Il primo consiste nella determinazione di un obiettivo. L'azione viene quindi prodotta tramite tre passaggi: determinazione dell'intenzione, pianificazione della sequenza di azioni e, infine, esecuzione. Il risultato è poi valutato attraverso un'altra sequenza di passaggi complementari agli stadi di esecuzione: percezione dell'output, interpretazione, valutazione, cioè paragone fra stato raggiunto e stato concepito.

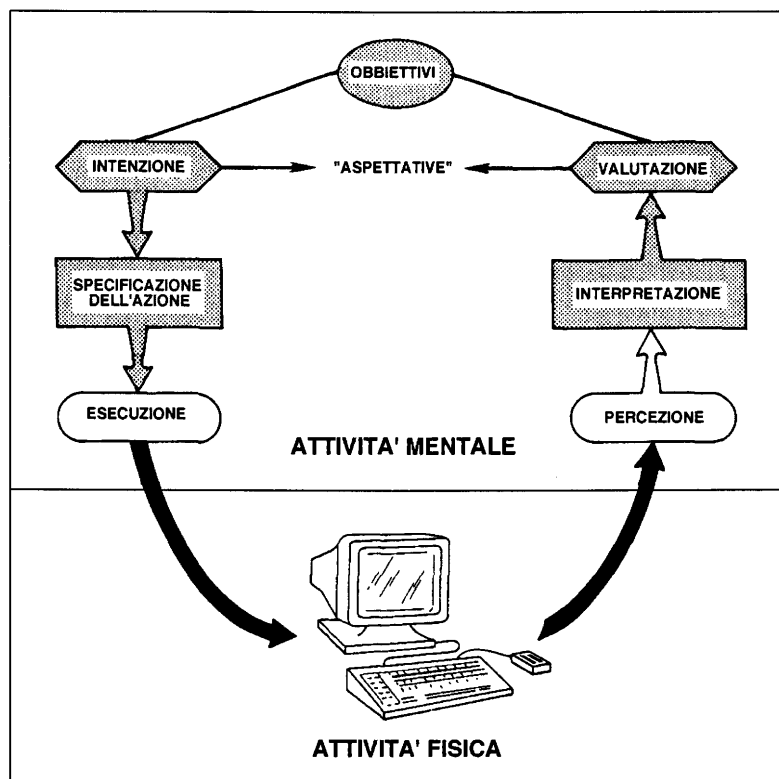


Figura 5. Modello dell'attività umana (adattato da Norman, 1986 pag. 42).

Secondo l'approccio globale a livello del primo stadio assumono grande importanza la *motivazioni* e *atteggiamenti*. A parità di scopo oggettivo variazioni di questi fattori possono influenzare anche largamente il comportamento di interazione

In questo contesto con il termine motivazione ci si riferisce all'aspetto dinamico dell'interazione, vale a dire al processo che guida il comportamento verso alcune categorie preferenziali di situazioni o oggetti (Nuttin, 1980). L'importanza della motivazione nell'interazione con sistemi flessibili è legata all'allargamento della situazione dialogica che permette di raggiungere uno stesso obiettivo mediante strategie comunicative e di

soluzione profondamente diverse. Tale concetto viene dunque considerato come una determinante fondamentale del comportamento dell'utente. Il suo effetto può essere controllato o manipolato in vari modi, alcuni dei quali sono già stati presentati (§ 3.4.1).

L'atteggiamento è un sentimento generale e durevole nei confronti di una certa persona od oggetto (Fishbein e Ajzen,). L'atteggiamento verso il computer è considerato una variabile di particolare rilievo al fine di prevedere la reazione psicologica e quindi il comportamento interattivo dell'utente (Reece, 1982; Wagman, 1983; Loyd e Gressard, 1984; Dambrot *et al.*, 1985; Popovich, *et al.*, 1987, Coover e Delcourt, 1992; Sensales e Greenfield, 1991). Attualmente sono disponibili numerose scale psicometriche che permettono di misurare tale costrutto (Loyd e Loyd, 1985; Dukes, *et al.*, 1989; Marcoulides, 1989; Kernan e Howard, 1990; Harrison e Rainer, 1992; Bandalos e Benson, 1990, Zabrajsek, Waters, Popovich, Craft e Hampton, 1990). Per quanto non sia ancora del tutto stabilita l'efficacia predittiva di tali strumenti rispetto all'accettazione del computer e all'apprendimento dei contenuti informatici, alcuni lavori hanno evidenziato l'esistenza di una correlazione fra accettazione e familiarità indotta dall'esperienza diretta (Dambrot *et al.* 1985, Loyd e Loyd, 1985; Gilroy e Desai, 1986; Szaina, 1994; Marcoulides, 1989; Gilroy e Desai, 1986; Cohen e Waugh, 1989; Parasumaran e Igarbia, 1990). Al crescere della familiarità tende a migliorare l'atteggiamento verso il computer (§5); anche la mera esposizione tende ad alleviare eventuali sentimenti di disagio (Cohen e Waugh, 1989). E' dunque necessario analizzare più nel dettaglio i principi di unificazione formale che legano l'atteggiamento alle altre determinanti dell'interazione. Si può infatti prevedere che l'atteggiamento diventi tanto più predittivo quanto più il computer è percepito come un artefatto socio-cognitivo.

A livello di esecuzione e valutazione, oltre alle capacità cognitive dell'utente, assume un'importanza fondamentale la sua *esperienza* passata (Nielsen, 1993; Marquière, 1994). Il ruolo dell'esperienza sul significato del nostro mondo percettivo è comunemente accettato: il senso di familiarità che il computer ha per l'utente e il particolare significato che esso assume nel suo mondo privato sono il risultato del numero e del tipo di esperienze che con il computer l'utente ha avuto. L'esperienza passata influenza l'interazione determinando la disponibilità di modelli per il comportamento. Se un utente ha avuto unicamente esperienza di sistemi altamente vincolati tenderà a riprodurre questo tipo di interazione anche con sistemi più flessibili (§ 7). Se l'utente ha avuto in passato esperienze fallimentari con sistemi informatici, il suo atteggiamento e la sua motivazione sicuramente ne risentiranno.

Ai fini di prevedere e comprendere la prestazione dell'utente è necessario distinguere fra tre dimensioni di esperienza (Nielsen, 1993): (a) esperienza informatica generale; (b) esperienza con il particolare sistema oggetto di studio; (c) esperienza con il compito. Le

ultime due dimensioni sono già state discusse relativamente alle determinanti del compito, occupiamoci dunque della prima.

L'esperienza informatica generale è tipicamente più difficile da misurare e concettualizzare. A lungo è stata identificata con l'*alfabetizzazione informatica* (*computer literacy*; van Vliet, Kletke e Chakraborty, 1994), vale a dire con l'insieme di conoscenze e abilità sottostanti l'uso del computer. Questo è un concetto controverso che si è andato via via modificando insieme allo sviluppo delle tecnologie informatiche. Inizialmente venne associato all'abilità di programmazione, ma, quando il graduale processo di semplificazione delle interfacce rese l'elaboratore un mezzo per l'esecuzione di compiti non specificamente informatici, fu esteso alla capacità di utilizzare specifici pacchetti di software (§ 1.2), o più generalmente alla capacità di agire in una società moderna (Rosenberg, 1991). Più recentemente, invece è stato definito come l'abilità di usare il computer a scopi diversi, mostrando un elevato controllo sull'interazione (van Vliet et al, 1994). Sulla base di quest'ultima definizione sono stati sviluppati uno strumento di misurazione oggettiva e uno di autovalutazione. Entrambi mostrano vantaggi e svantaggi, ma il problema principale consiste nella mancanza di correlazione fra le due misure. In generale, l'autovalutazione sembra indurre fenomeni di sovrastima, specialmente fra utenti inesperti (van Vliet, 1994).

Le difficoltà legate alla concettualizzazione e alla misurazione dell'alfabetizzazione informatica hanno fatto sì che tale determinante venga spesso sottovalutata, se non addirittura ignorata nelle valutazioni. L'approccio globale propone dunque di sostituire a tale costrutto quello di familiarità o esperienza. Piuttosto che chiedere a un utente di valutare la sua abilità nell'uso di un determinato programma, conviene chiedere quanto spesso lo utilizza e il suo livello di soddisfazione medio. Tali indici, per quanto più indiretti, sono meno valutativi e dovrebbero quindi fornire risultati più affidabili (§ 5). Inoltre, dato il generale processo di semplificazione dell'interazione, i due concetti stanno effettivamente diventando sinonimi.

Riassumendo, secondo l'approccio globale l'utente influenza l'interazione tramite varie determinanti. Il suo comportamento è mediato da tre agenti che contribuiscono in modo correlato all'azione. Il primo è l'*agente cognitivo puro*, l'elaboratore di informazioni secondo l'approccio cognitivista, caratterizzato da precisi limiti strutturali e procedurali che devono essere considerati durante l'analisi. Il secondo è l'*agente socio-cognitivo*, un processore di informazioni il cui comportamento è vincolato dal modo di percepire il computer, il compito e l'ambiente esterno, da atteggiamenti interni e dal contesto sociale (§ 4.2.1). Il terzo è l'*agente individuale*, l'utente che si pone di fronte al computer con specifiche motivazioni e esperienze. In questa prospettiva, dunque, il comportamento umano viene analizzato da un punto di vista assai più ampio rispetto alla tradizione ergonomica.

4.4. *Analisi dei risultati*

Secondo l'approccio globale la fase di analisi dei risultati serve a fornire un quadro completo dell'usabilità del sistema informatico. È necessario, dunque, tenere in considerazione tutte le possibili determinanti dell'interazione che possono influenzare il valore dei singoli attributi di usabilità.

Nel terzo capitolo (§ 3.1) sono stati proposti due modelli del concetto di usabilità che si differenziano sul numero e sulla tipologia degli attributi ritenuti salienti. Entrambi i modelli però descrivono l'usabilità come un concetto misto, composto da alcuni attributi oggettivamente misurabili e da altri che derivano invece dal giudizio dell'utente. Storicamente le misure oggettive sono state privilegiate ma, negli anni, la crescente importanza attribuita al fattore umano ha modificato tale tendenza (Ives e Olson, 1984; Baroudi, Olson e Ives, 1986; Login, Diehl e Norman, 1988).

Comunque il problema della correlazione fra misure soggettive e misure oggettive è stato a lungo ignorato (Muckler, 1992). Eppure tale aspetto rappresenta un elemento chiave per la comprensione del sistema utente-computer e per la verifica della validità della valutazione. Poiché la maggior parte dei lavori che hanno affrontato l'argomento si riferiscono a valutazioni di sistemi di *Information Retrieval* (IR), apriamo una breve parentesi per spiegare come viene solitamente operazionalizzata la loro usabilità.

I sistemi di IR sono strumenti informatici capaci di immagazzinare, organizzare e recuperare ingenti quantità di informazione (Salton e McGill, 1983). L'aspetto valutativo è sempre stato considerato molto importante per il loro sviluppo, probabilmente grazie all'abitudine al lavoro interdisciplinare delle Scienze dell'Informazione e al fatto che da sempre tali sistemi si rivolgono a un'utenza eterogenea. Eppure, a prova della complessità del lavoro di valutazione, anche in questo settore tuttora si sente la mancanza di affidabili misure e di adeguate metodologie (Robertson e Hancock, 1992; Salton, 1992; Su, 1992; Tague-Sutcliffe, 1992). Gli attributi di usabilità di un sistema di IR sono solitamente definiti sulla base della proposta degli Standard ISO. L'*efficacia* coincide con la *rilevanza* dell'informazione recuperata. Tale indice viene solitamente ottenuto combinando due misure proporzionali (Tague-Sutcliffe, 1992): una di completezza (*recall*) e una di precisione (*precision*)³². L'*efficienza* riguarda il tempo impiegato per il recupero dell'informazione e il costo economico dell'operazione. La *soddisfazione dell'utente* viene misurata da due indici distinti: uno generale, riguardante il giudizio sul sistema; uno

³² *Recall* = rr/rt , dove rr corrisponde al numero di documenti rilevanti recuperati e rt al numero di documenti rilevanti presenti nel data-base. *Precision* = rr/tr , dove tr corrisponde al numero totale dei documenti recuperati. L'uso di queste misure ha vari limiti. La difficoltà principale riguarda la determinazione non ambigua dell'insieme dei documenti rilevanti, compito particolarmente gravoso nel caso di data-base estesi (Su, 1992; Sutcliffe, 1992).

specifico, riguardante la *soddisfazione verso l'informazione* ottenuta (Ives, Olson e Baroudi, 1983).

La valutazione completa di un sistema di IR implica l'utilizzo di tutte queste misure (Su, 1992; Tague-Sutcliffe, 1992). Tuttavia sono state evidenziate correlazioni negative fra *recall* e *precision* (Cleverdon, 1972), soddisfazione dell'utente e *precision* (Hitchingam, 1979), giudizio sull'utilità di una ricerca e numero di documenti utili effettivamente recuperati, nonché basse correlazioni fra *precision* e *willingness to pay*³³ (Tagliacozzo, 1977). Il primo risultato è l'unico facilmente spiegabile: quanto più si estende una ricerca, per aumentare il numero di documenti rilevanti, tanto meno preciso sarà il risultato finale. Tutti gli altri sono invece controintuitivi e sembrano suggerire che misure oggettive e misure soggettive forse non sono compatibili.

La mancanza di correlazione fra misure oggettive e soggettive può essere attribuita a due cause antitetiche: (a) effettiva disomogeneità fra le misure; (b) errori nella misurazione. Se accettiamo la prima possibilità siamo costretti a concludere che l'utente non è un buon giudice dell'interazione (Nielsen, 1993). Lo sviluppo di questa idea conduce però a un paradosso: l'approccio centrato sull'utente è stato sviluppato per agevolare il fattore umano realizzando sistemi informatici adeguati alla sua natura. L'utente, comunque, non sembra in grado di discriminare l'usabilità e per qualche strano motivo potrebbe privilegiare sistemi poco usabili e respingere sistemi eccellenti. Di conseguenza l'approccio centrato sull'utente perde valore. Personalmente, sono poco propensa ad accettare questa conclusione. Ritengo piuttosto che le attuali metodologie di valutazione siano inadeguate a comprendere le variabili che effettivamente influenzano il giudizio dell'utente. Infatti, oltre che da caratteristiche del sistema, la soddisfazione potrebbe dipendere anche da caratteristiche dell'utente che vanno tenute in considerazione per evitare indesiderati effetti di mascheramento.

³³ Tale variabile rappresenta una stima della spesa che un utente sarebbe disposto a sostenere per l'utilizzo o l'acquisto di un dato sistema.

4.5. Conclusioni

Siamo giunti alla fine di un lungo percorso teorico in cui sono stati presentati numerosi concetti innovativi e certamente problematici. I principali riguardano l'estensione del concetto di computer a artefatto socio-cognitivo e la proposta di un'analisi empirica che tenga in considerazione un ampio numero di variabili e che produca teorie di campo. Prima di rifiutare o accettare la proposta è utile vederla applicata. A questo scopo è dedicata la seconda parte del lavoro che si concentra particolarmente sulla comprensione del ruolo dell'esperienza nell'interazione con sistemi interattivi flessibili.

Nel prossimo capitolo vengono riportati i risultati di due studi correlazionali che rappresentano un contributo alla categorizzazione dell'utente, fornendo nuovi strumenti di misurazione e un modello dei rapporti fra le sue determinanti. In seguito vengono descritte due valutazioni di sistemi multimodali in cui la pragmatica dell'approccio viene applicata a esperimenti di simulazione.

Parte seconda:
Applicazioni del modello

5. Le determinanti relative all'utente

La principale difficoltà sottostante l'applicazione dell'approccio globale è la gestione dell'elevato numero di variabili considerate determinanti nell'interazione. L'entità in assoluto più complessa è l'utente. Prima di applicare il metodo a reali situazioni interattive è dunque necessario semplificare la rappresentazione di tale entità, raccogliendo informazioni sull'organizzazione sottostante le sue determinanti. Questa fase rappresenta in qualche modo uno studio pilota, che permette di tarare gli strumenti di misura e di elaborare una prima rappresentazione dell'utente, che verrà poi sottoposta a verifica in situazioni di interazione.

I due studi presentati in questo capitolo sono dedicati alla comprensione delle relazioni fra alcune determinanti il cui effetto sull'interazione e le cui reciproche influenze sono maggiormente controversi. Questa parte della tesi affronta esclusivamente la seconda problematica, cercando di proporre l'organizzazione interna delle determinanti. In particolare vengono analizzati i rapporti fra: esperienza informatica, atteggiamento verso il computer, sesso ed età. Ricordiamo in breve i principali risultati riportati dalla letteratura (§ 4.3) sottolineando per ognuno di essi il contributo dei lavori presentati.

La conoscenza informatica è comunemente considerata importante ai fini di predire l'interazione, ma sono state evidenziate svariate difficoltà legate alla sua operazionalizzazione. La principale è un generalizzato *bias* di sovrastima, per cui in situazioni di autovalutazione le persone tendono a riportare una conoscenza maggiore di quella effettivamente posseduta (§ 4.3.1). Gli studi si propongono di sostituire al concetto di *conoscenza* quello più facilmente misurabile di *esperienza* e di verificare la loro intercambiabilità. A questo fine viene elaborato e validato un questionario. Un ulteriore scopo è l'identificazione delle variabili che meglio permettono di prevedere l'esperienza. La letteratura ne evidenzia tre, sesso, età e atteggiamento, ma il loro valore predittivo è tuttora oggetto di discussione. Nelle valutazioni di usabilità queste conoscenze sono estremamente utili, poiché consentono di verificare l'affidabilità delle autovalutazioni o di inferire il livello di conoscenza nel caso in cui non sia possibile o consigliabile misurarlo direttamente.

L'influenza del sesso è spiegata in modo diverso dall'ipotesi disposizionalista e situazionalista (§ 4.3). Entrambe le spiegazioni partono dalla constatazione che le donne hanno meno esperienza degli uomini, ma si differenziano nell'identificazione delle cause sottostanti tale divario. Le principali differenze sono riassunte dalla Tabella 1.

Iph. disposizionalista	Iph. situazionalista
Atteggiamento _M > Atteggiamento _F	Atteggiamento _M = Atteggiamento _F
Programmazione _M > Programmazione _F Videoscrittura _M > Videoscrittura _F	Videogiochi _M > Videogiochi _F
Volontà _M > Volontà _F	Disponibilità _M > Disponibilità _F

Tabella 1. Principali differenze fra ipotesi disposizionalista e situazionalista.

Le differenze ipotizzate riguardano tutte il sesso biologico di una persona. Un modo alternativo di testare la validità delle due ipotesi implica di considerarne l'identità sessuale, vale a dire il modo in cui gli individui, indipendentemente dal loro effettivo sesso biologico, si percepiscono sulla base delle dimensioni sessuali (Ashmore e Del Boca, 1981). A questo livello si possono distinguere quattro tipologie di persone: maschili, femminili, androgine e indifferenziate (Bem, 1978 e 1981). Le prime si descrivono prevalentemente sulla base di caratteristiche proprie del sesso maschile, le seconde di quello femminile. Le terze invece tendono a identificarsi con intensità simile con le caratteristiche di entrambi i sessi, mentre le quarte non si identificano nè con le caratteristiche di un sesso nè con quelle dell'altro.

Sulla base dell'identità sessuale possiamo evidenziare un'ulteriore differenza fra ipotesi situazionalista e ipotesi disposizionalista. Secondo la prima il divario fra donne e uomini è dovuto a intrinseche caratteristiche dei due sessi. Esperienza e atteggiamento dovrebbero dunque essere modulate secondo il seguente ordinamento: $IS_m > (IS_a \& IS_i) > IS_f$. Secondo l'ipotesi situazionalista, le caratteristiche dei due sessi sono ininfluenti e di conseguenza dovrebbe esserlo anche l'identità sessuale.

Le ricerche qui presentate si propongono di confermare l'ipotesi situazionalista.

5.1. Studio 1

Questo lavoro è parte di un progetto di ricerca più ampio finalizzato alla valutazione dell'andamento del processo di alfabetizzazione informatica nella realtà universitaria italiana, che prevede una seconda raccolta di dati su un campione omogeneo a distanza di cinque anni. La prima fase, eseguita nella primavera del 1992, era subordinata a tre obiettivi principali:

1. Costruzione e validazione di un questionario per la misurazione dell'esperienza informatica; taratura su un campione italiano di un questionario inglese per la misurazione dell'atteggiamento verso il computer (Dambrot, *et al.*, 1985).
2. Verifica dell'ipotesi situazionalista.
3. Determinazione dei fattori che influenzano l'esperienza.

5.1.1. Metodo

5.1.1.1. Popolazione della ricerca

Il set di questionari è stato somministrato a un campione di 270 studenti dell'Università degli Studi di Trieste. Al momento della somministrazione, l'età media dei partecipanti era di circa 21 anni, da un minimo di 18 a un massimo di 40. Il 78% degli studenti frequentava il biennio, mentre solo il 4% erano fuori corso. Il 70% circa del campione è costituito da studenti di psicologia, il 20% da studenti di biologia, e il rimanente 10% è suddiviso fra Lettere e Filosofia, Economia e Commercio, Giurisprudenza. Volutamente non sono stati intervistati studenti di facoltà tecniche in cui l'informatica rappresenta una fondamentale materia di studio e che solitamente mostrano di valutare il computer più positivamente dei loro colleghi di altre discipline (Sensales e Greenfield, 1991).

5.1.1.2. Strumenti

La ricerca prevedeva la compilazione di tre questionari, in ordine di somministrazione:

1. Questionario *EC* (per la rilevazione dell'Esperienza con il Computer)
2. Questionario *CATT* (per la rilevazione dell'atteggiamento verso il computer)

3. Questionario *BSRI* (per la rilevazione del livello di androginia psicologica)

Questionario *EC*

La costruzione dello strumento è basata sulle indicazioni disponibili in letteratura (§ 4.3) e sui risultati di una serie di brevi studi pilota, ponendo grande attenzione al controllo del *bias* di sovrastima (§ 4.3.1). Il concetto di esperienza è stato dunque concettualizzato come familiarità indotta dalla frequenza d'uso, piuttosto che come insieme di abilità e conoscenze. Inoltre, in tutto lo strumento vengono rigorosamente evitati riferimenti valutativi e proposte domande incrociate nonché numerose richieste di specificazione (vedi Appendice 1).

Il questionario inizia con una breve sezione che permette di raccogliere alcuni fondamentali dati socio-anagrafici, quali età, sesso, corso di laurea, anno di corso, diploma di maturità ed un'eventuale qualifica professionale. Prosegue poi con 21 domande di cui una parte è finalizzate alla raccolta di auto-descrizioni delle proprie esperienze e l'altra alla verifica oggettiva delle conoscenze dichiarate.

La frequenza d'uso viene misurata su scala Likert a 5 categorie di risposta, da **1 mai** a **5 molto spesso**. La prima domanda (*hai mai utilizzato un computer?*) ha funzione di filtro. Nel caso in cui l'intervistato risponda *mai* invita a passare direttamente alla domanda 21, altrimenti a proseguire la compilazione secondo l'ordine previsto. I sei item successivi si riferiscono alla frequenza d'uso di alcune applicazioni informatiche (nell'ordine: posta elettronica, programmazione, video giochi, analisi di dati, videoscrittura e Information Retrieval). La domanda 8 ha lo scopo di misurare la frequenza di lettura di riviste informatiche, mentre la 9 la conoscenza di linguaggi di programmazione. Per ridurre il bias di sovrastima ogni item funge da filtro a una richiesta di specificazione. Se l'intervistato risponde di avere utilizzato almeno *raramente* una particolare applicazione, viene invitato a specificare il nome del software o il posto dove lo aveva utilizzato. Tale strategia permette di riconoscere già in fase di immissione dei dati eventuali incongruenze che comunque, come osservato durante uno studio pilota, vengono notevolmente diminuite. La domanda 10 verifica la volontà di accrescere la propria conoscenza ed infine l'undicesima, l'eventuale frequenza di corso informatici. Le successive richiede nell'ordine di dichiarare dove hanno imparato a usare il computer e se ne possiedono uno.

L'item 14 funge da filtro per la parte di verifica oggettiva, discriminando fra utenti di ambiente operativo MS-DOS³⁴, Macintosh o di entrambi. A seconda della risposta gli intervistati sono invitati a rispondere a 3 domande per ogni sistema operativo, articolate su una scala di difficoltà e selezionate sulla base di un breve studio pilota in cui sono stati

³⁴ Si noti che all'epoca della raccolta Windows 95 non era ancora disponibile e le due versioni di Windows 3.0-3.1 non erano molto diffuse.

coinvolti 10 esperti informatici. Il questionario si chiude con un item volto a verificare la robustezza dello stereotipo che descrive gli uomini come più bravi delle donne nell'utilizzare il computer.

Questionario CATT

Il CATT è stato sviluppato per misurare l'atteggiamento verso il computer in una popolazione di studenti universitari americani. Il questionario è formato da 20 item, 9 affermazioni positive e 11 negative sul computer. Ogni frase richiede un giudizio su una scala di tipo Likert a 5 punti, dove **5** significa *completo accordo* e **1** *completo disaccordo*. L'indice complessivo di atteggiamento viene ricavato mediando i punteggi dei singoli item, previo ribaltamento delle affermazioni positive. La versione italiana della scala è stata tradotta da due persone curando la massima fedeltà linguistica (vedi Appendice 1). Successivamente è stata condotta una traduzione *backward* da parte di un'interprete di madre lingua italiana.

La scelta del questionario CATT è dovuta alla sua provata affidabilità interna, esterna e di costrutto. Originariamente gli autori sottoposero la scala a due tarature che fornirono valori di attendibilità soddisfacenti, nell'ordine $\alpha = .84$ e $\alpha = .79$ (Dambrot *et al.* 1985). Risultati simili sono emersi anche in ricerche condotte da altri autori (Ogletree e Williams, Kernan e Howard, 1990), che hanno inoltre evidenziato elevate correlazioni con altre quattro scale³⁵ che misurano lo stesso costrutto (Zabrajsek, Waters, Popovich, Craft, Hampton, 1990).

Questionario BSRI

Nelle intenzioni dell'autrice Sandra Bem (1974, 1981) tale test serve alla misurazione di un tratto di androginia psicologica. Il reattivo costituisce un pilastro del progresso epistemico e metodologico dello studio dell'identità sessuale, in quanto si basa sul presupposto che in un medesimo individuo possano coesistere componenti maschili e femminili. Il BSRI fornisce infatti due punteggi separati, uno di mascolinità (scala M), l'altro di femminilità (scala F) che non sono concettualmente opposti ma completamente privi di correlazione. Inoltre, le due scale misurano concetti normativi e non derivati da giudizi di frequenza.

Il BSRI è composto da 60 item, ciascuno dei quali costituito da una breve frase o da un aggettivo designante caratteristiche di personalità. Di questi, 20 esprimono caratteristiche tipicamente maschili, 20 stereotipicamente femminili e gli altri hanno funzione di riempitivi (Bem, 1981). La versione utilizzata nello studio corrisponde alla taratura italiana presentata nel 1986 da De Leo e colleghi (De Leo, Villa, Magni,

³⁵ ZOL12, CARS, NCAS, ATTCUS, in tutti i casi l'indice di correlazione è risultato superiore a .70.

Andreatta, Gagliardi, 1986). Tale versione ha fornito risultati soddisfacenti a tutte le prove psicometriche cui è stata sottoposta: stima dell'attendibilità come consistenza interna (coefficiente α di Cronbach), come stabilità (procedura test-retest) e validazione strutturale (verifica dell'incorrelazione tra le due scale).

5.1.1.3. *Procedura*

La somministrazione dei questionari avveniva in gruppi di massimo trenta persone, sotto la supervisione di due sperimentatori esperti che avevano il compito di registrare e chiarire eventuali dubbi di interpretazione. Prima della somministrazione ai partecipanti veniva garantita l'assoluta anonimità e l'ininfluenza delle loro risposte ai fini della valutazione accademica. I questionari EC e CATT venivano consegnati insieme, mentre il BSRI veniva presentato come una ricerca autonoma. La compilazione dell'intero set di questionari richiedeva al massimo 25 minuti.

5.1.2. **Risultati**

5.1.2.1. *Taratura degli strumenti*

Questionario EC

Per verificare la consistenza interna delle domande di autovalutazione dell'esperienza sono state eseguite una serie di correlazioni fra item o indici incrociati. Per garantire la massima validità, il campione è stato sottoposto a rigorosi controlli che hanno permesso di escludere tutti gli intervistati, la cui inconsistenza nelle risposte suggeriva la presenza del *bias* di sovrastima.

Innanzitutto è stata verificata la congruenza fra il *Giudizio Diretto sull'Esperienza* (item 1; GDE) e una stima di tale valore ottenuta mediando le risposte agli item specifici (2-7; *Giudizio Indiretto Esperienza* GIE). La correlazione è risulta soddisfacente, $r = .82$, ($n = 202$), $p < .001$. Successivamente GIE è stato trasformato in un indice a quattro categorie, corrispondenti ai quartili della sua distribuzione, e inserito in una tabella di contingenza insieme ai valori di GDE. Ciò ha permesso di evidenziare tre casi chiaramente incongruenti, che sono stati eliminati dal campione.

Un'ulteriore verifica ha riguardato il paragone fra gli item volti a misurare esperienza e conoscenza della programmazione (domande 3 e 9). Nuovamente si evidenzia un'elevata correlazione, $r = .78$, ($n = 201$) $p < .001$. Il test di Wilcoxon dimostra che il lieve scarto fra le due variabili è dovuto a una significativa sovrastima dell'item 3, che non era seguito da richiesta di specificazione, $Z = -1.68$, ($n = 80$), 1-tail p

< .05. Tutti i casi in cui lo scarto fra esperienza e conoscenza era superiore a 1 sono stati esclusi ($n = 3$). L'analisi incrociata di GDE e di una delle alternative dell'item 12 ha permesso di evidenziare ed eliminare altri due casi incongruenti, che pur avendo dichiarato di usare il computer frequentemente ammettevano poi di non saperlo usare autonomamente.

L'elevata correlazione fra GDE e GIE, la migliore distribuzione statistica di GDE e la prevalente tendenza a utilizzare un'unica applicazione hanno indotto a preferire GDE su GIE come *misura generale di esperienza*. La distribuzione di GDE è stata quindi normalizzata tramite calcolo della radice quadrata \sqrt{GDE} . Per ogni partecipante è stato invece calcolato un *indice oggettivo di conoscenza* (IOC) attribuendo alle domande di test un valore da 1 a 3 a seconda della loro difficoltà e sommando il punteggio così ottenuto. L'elevata correlazione fra GDE e IOC dimostra che il singolo item è effettivamente in grado di discriminare i partecipanti sulla dimensione desiderata, $r = .61$, ($n = 79$), $p < .001$. Si noti però, che il limitato numero di casi su cui l'analisi è stata eseguita è dovuto a un notevole numero di *missing-value* nelle risposte ai test oggettivi.

Per ogni partecipante è stato inoltre costruito un *indice di Uso* a 4 categorie di risposta, (nullo, raro, medio e frequente) da utilizzare come variabile indipendente. Alla prima categoria appartengono tutti i partecipanti che hanno dichiarato di non avere mai usato il computer. Le altre corrispondono invece ai terzili della media ottenuta dai valori di GDE e delle applicazioni più frequentemente usate (item 3, 4, 6 e 9).

Questionario CATT

La scala per la misurazione *dell'atteggiamento verso il computer* è stata sottoposta ad un'analisi di attendibilità per misurarne la coesione interna, le correlazioni item-totale e l'affidabilità nel campione in esame. L'attendibilità della scala è risultata inferiore ai valori originariamente trovati da Dambrot e colleghi ($\alpha = .77$). L'analisi congiunta dei valori di correlazione item-totale, di correlazione multipla quadrata e del guadagno in termini di attendibilità suggerisce di escludere cinque item (10, 11, 12, 13 e 20) in quanto poco correlati al resto della scala ($r < .25$ e $r^2 < .18$). In questo modo l'attendibilità raggiunge un valore soddisfacente ($\alpha = .83$).

Per analizzare più nel dettaglio la struttura dello strumento, i punteggi relativi ai 20 item originali sono stati sottoposti ad analisi fattoriale mediante il metodo degli assi principali. I due fattori estratti tramite rotazione VARIMAX soddisfano i criteri dello *screen-test* e spiegano il 33.5% della varianza (il primo ha un autovalore di 4.45 ed è responsabile del 22.3% della varianza, il secondo invece ha un autovalore di 2.23). Per circoscrivere il significato dei fattori è stato adottato il seguente criterio di selezione degli item: vengono esclusi tutti gli item spuri (entrambe le saturazioni $< .40$), e misti (una

saturazione > .30 e l'altra > .40). La matrice delle saturazioni fattoriali è riportata nella Tabella 2.

Il primo fattore (che satura gli item 4, 6, 8, 9, 18 e 19) rappresenta una dimensione emotiva di giudizio che ha all'estremo negativo il vissuto di minaccia, dovuto all'induzione di timore per l'invasione e la sottrazione di controllo sulle nostre vite da parte del computer, mentre all'estremo positivo il vissuto di beneficio e sicurezza, per cui la macchina è vissuta come rassicurante. Il secondo (che satura gli item 1, 2, 7, 14, 15 e 16) rappresenta invece una dimensione più legata allo scambio interattivo ai cui poli estremi si situano i concetti di ostilità e amichevolezza.

Item	Fattore I Rassicurazione	Fattore II Amichevolezza
Item 1	0,31208	0,55103
Item 2	0,02746	0,46714
Item 3	0,57157	0,43027
Item 4	0,41962	0,04844
Item 5	0,36613	0,22979
Item 6	0,41258	0,08043
Item 7	0,17036	0,69342
Item 8	0,73246	-0,06074
Item 9	0,73062	-0,08329
Item 10	0,31564	-0,21256
Item 11	0,21227	0,09734
Item 12	0,00949	-0,16082
Item 13	0,13417	0,09157
Item 14	0,10276	0,50314
Item 15	0,11521	0,56251
Item 16	0,29600	0,59313
Item 17	0,05831	0,36229
Item 18	0,54134	0,27850
Item 19	0,54341	0,19645
Item 20	0,01219	0,20619

Tabella 2. Taratura del CATT (20 item).

La tabella riporta la matrice delle saturazioni fattoriali. I numeri in grassetto sulle colonne dei fattori indicano le saturazioni fattoriali ritenute soddisfacenti; nella prima colonna gli item selezionati in base al criterio descritto nel testo.

I punteggi dei singoli item sono stati mediati in due indici riassuntivi congruentemente con la configurazione fattoriale. Poiché in entrambe i casi i giudizi si situano nel lato positivo della dimensione gli indici sono stati rispettivamente etichettati come *rassicurazione* e *amichevolezza*. Inoltre è stata calcolata una misura di *atteggiamento complessivo*, mediando il valore delle risposte ai 15 item della nuova scala. I valori medi

di tali indici e i risultati di tre separati t-test che ne valutano la differenza dal punto neutro della scala (3) sono riportati nella Tabella 3.

Indice	Media	SD	<i>t</i>	gl	2-tail p
Generale	3.52	.65	12.71	253	.000
Amichevolezza	3.43	.83	8.35	262	.000
Rassicurazione	4.04	.77	- 21.86	262	.000

Tabella 3. Descrittive dei tre indici di atteggiamento verso il computer.

I valori di *t* si riferiscono al paragone fra le medie e il valore neutro della scala (3).

Tutti gli item sono stati ricodificati in modo tale che 1 indichi l'estremo negativo e 5 l'estremo positivo dell'atteggiamento. Alla variabile assicurazione è stata quindi applicata una trasformazione logaritmica (\log_{10} assicurazione), che ha permesso di diminuire la skewness positiva, il numero di outlier e migliorare la normalità.

Questionario BSRI

I dati ricavati dall'applicazione del reattivo sono stati sottoposti alle procedure di elaborazione suggerite da Bem nel *Professional Manual* (1981). Entrambe le scale sono state sottoposte ad analisi di attendibilità, che ha fornito risultati migliori per la scala M ($\alpha = .88$) rispetto alla scala F ($\alpha = .75$). Poiché lo studio della validità fattoriale dello strumento e della consistenza interna delle sottoscale trascendeva gli obiettivi dello studio, si è preferito ignorare la differenza e procedere alla costruzione degli indici riassuntivi rispettando la configurazione originale. Di conseguenza, i due indici M e F sono stati calcolati mediando tutti i valori dei singoli item delle due scale. In seguito, ogni individuo è stato classificato come *femminile*, *maschile*, *androgino* o *indifferenziato* sulla base della posizione dei suoi due punteggi F e M, rispetto alle mediane delle distribuzioni, calcolate per ogni sottoscala sull'intero campione. In generale, i soggetti si distribuiscono abbastanza uniformemente fra indifferenziati (27.5%), femminili (25%) e maschili (27.5%), mentre gli androgini risultano lievemente inferiori (20%).

5.1.2.2. Verifica dell'ipotesi situazionalista

Il quadro generale relativo all'alfabetizzazione informatica è sconcertante. Alla domanda diretta sull'esperienza (GDE), il 24% del campione dichiara di non avere *mai* utilizzato un computer. Se si aggiungono coloro che hanno risposto di utilizzarlo *raramente* si supera la metà dei partecipanti (57%). Solo il 22% dichiara di usarlo *a volte*,

il 12% *frequentemente* e il 9% *molto spesso*. Inoltre è interessante notare che il 45% degli utenti ammette di non sapere usare il computer autonomamente. Solo il 21% degli utenti ha frequentato un corso di computer, il 16% invece lo ha iniziato, ma non portato a termine, e il rimanente 63% non lo ha mai iniziato. La partecipazione a corsi informatici non risulta influenzata dal Sesso Biologico, come invece ci si sarebbe dovuti aspettare secondo l'ipotesi situazionalista.

A sostegno di tale ipotesi va invece la differenza significativa fra maschi e femmine nel numero di persone che hanno dichiarato di possedere un computer, $\chi^2(2) = 10.8$, $p < .01$. Le percentuali sono riportate in Tabella 4. Un'ANOVA su \sqrt{GDE} con Possesso (2) come fattore *between-subjects* indica inoltre che la disponibilità del computer influenza l'esperienza $F(1, 196) = 38.02$, $p < .001$. Coloro che lo possiedono dichiarano in media un'esperienza di 3.47, mentre quelli che non lo possiedono personalmente di 2.57.

	Femmine		Maschi	
Si	(36.6)	30	52	(63.4)
No, ma	(65.4)	34	18	(34.6)
No	(51.4)	36	34	(48.6)
Totale		100	104	

Tabella 4. Disponibilità del computer in funzione del sesso biologico.

Legenda: Si = possiedo un computer personale; No, ma = non possiedo un computer, ma posso utilizzarne uno ogni volta che mi serve; No = non possiedo un computer. I numeri fra parentesi rappresentano le percentuali di riga, quelli in grassetto il numero di casi.

In generale gli intervistati dichiarano di volere accrescere la loro conoscenza [media = 3.59, SE = .09, t-test contro il punto neutro della scala: $t(196) = 3.6$, $p < .001$]. A differenza di quanto previsto dall'ipotesi disposizionalista, il desiderio non dipende dal Sesso Biologico. Si noti però che nel campione dei partecipanti maschi al crescere dell'esperienza aumenta anche il desiderio di migliorare la propria conoscenza [$r = .39$, ($n = 104$), $p < .001$], mentre nel campione delle femmine le due variabili appaiono incorrelate [$r = .13$, ($n = 101$), $p = n.s.$].

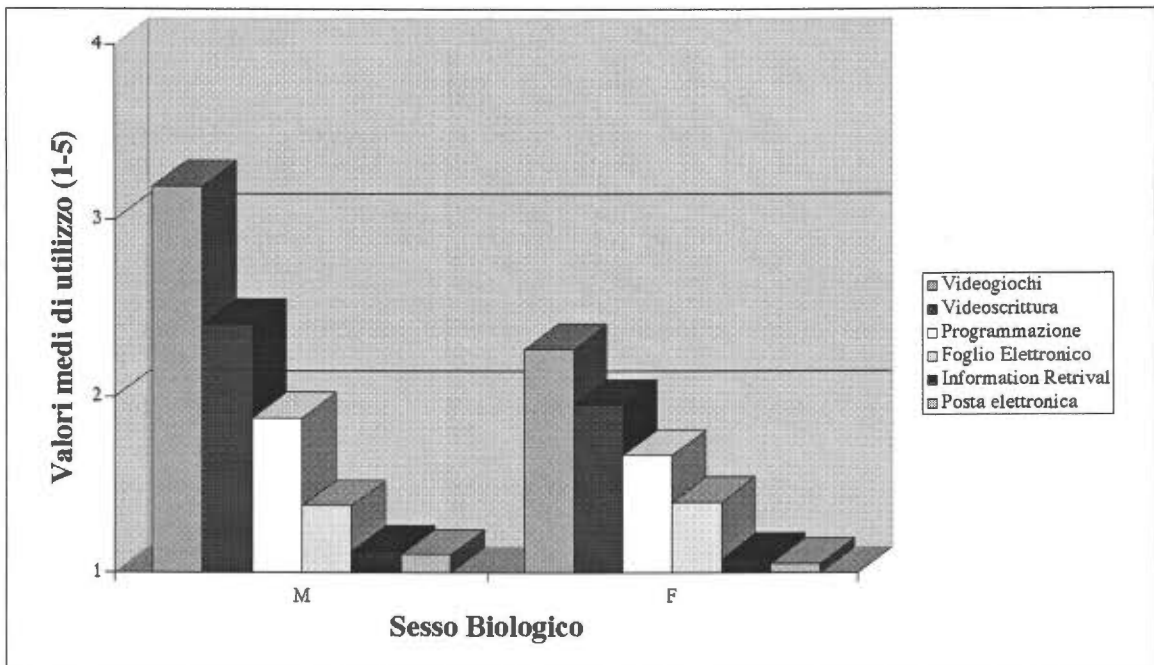


Figura 6. Utilizzo di applicazioni informatiche in funzione del Sesso Biologico.

Il grafico riporta i valori medi della frequenza di utilizzo dichiarata da 205 intervistati, cioè da quella parte del campione che aveva dichiarato di utilizzare il computer almeno raramente. La scala di frequenza si articola da 1 (mai) a 5 (molto spesso).

Come si può vedere dal grafico in Figura 6, il computer viene prevalentemente sfruttato come strumento ludico. Il generale squilibrio a favore dei maschi raggiunge la significatività solo su alcune applicazioni.

Data la distribuzione non normale che caratterizza tutte le variabili (curtosi e *skewness* positive estremamente accentuate), i confronti sono stati eseguiti mediante *Mann-Withney U test*. In generale i maschi usano videogiochi e programmi di videoscrittura più spesso delle femmine, nell'ordine $U = 4206.5$ ($n = 205$), $p < .05$ e $U = 4330.5$ ($n = 204$), $p < .05$. Per quanto riguarda la programmazione la differenza non raggiunge la significatività, $U = 4601.5$, ($n = 204$), $p = .11$. Questi risultati vanno dunque a favore della spiegazione situazionalista.

L'indice generale di atteggiamento e \sqrt{GDE} sono stati analizzati da una MANOVA per il disegno Sesso Biologico (2) * Identità Sessuale (4). Dal confronto multivariato emerge un significativo effetto del Sesso Biologico $F(2, 221) = 6.43$, $p < .01$. I risultati a livello univariato mostrano però che tale effetto è prevalentemente dovuto all'esperienza, $F(1, 222) = 12.83$, $p < .001$, mentre tende a scomparire nell'atteggiamento, $F(1, 222) = 3.06$, $p = .08$ (vedi Figura 7). L'Identità Sessuale risulta ininfluenza, sia come effetto principale che nell'interazione. In generale dunque i risultati tendono nuovamente a privilegiare l'ipotesi situazionalista.

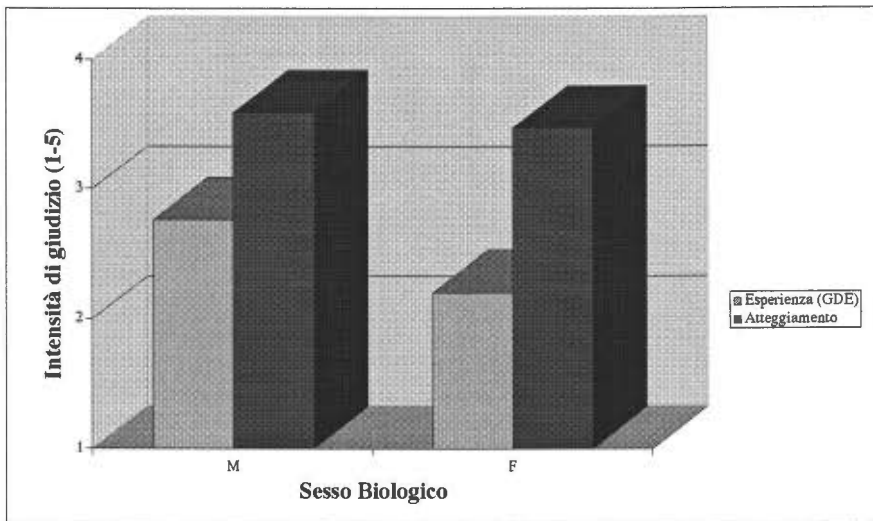


Figura 7. GDE e atteggiamento in funzione di Sesso Biologico e Identità Sexuale.

L'asse delle ordinate riporta i valori medi delle utovalutazioni relative alla propria esperienza informatica e all'atteggiamento nei confronti del computer.

Per analizzare più nel dettaglio l'atteggiamento nei confronti del computer amichevolezza e \log_{10} rassicurazione sono state sottoposte a una MANOVA per il disegno Sesso Biologico (2) * Uso (4). L'effetto della frequenza d'uso è evidente $F(6, 478) = 7.41, p < .001$. Al crescere dell'esperienza aumenta l'amichevolezza e la rassicurazione indotta dal computer (entrambe le univariate sono significative a $p < .001$). Il Sesso Biologico non dimostra alcuna influenza né a livello di effetto principale né nell'interazione, come si evidenzia dal grafico in Figura 8.

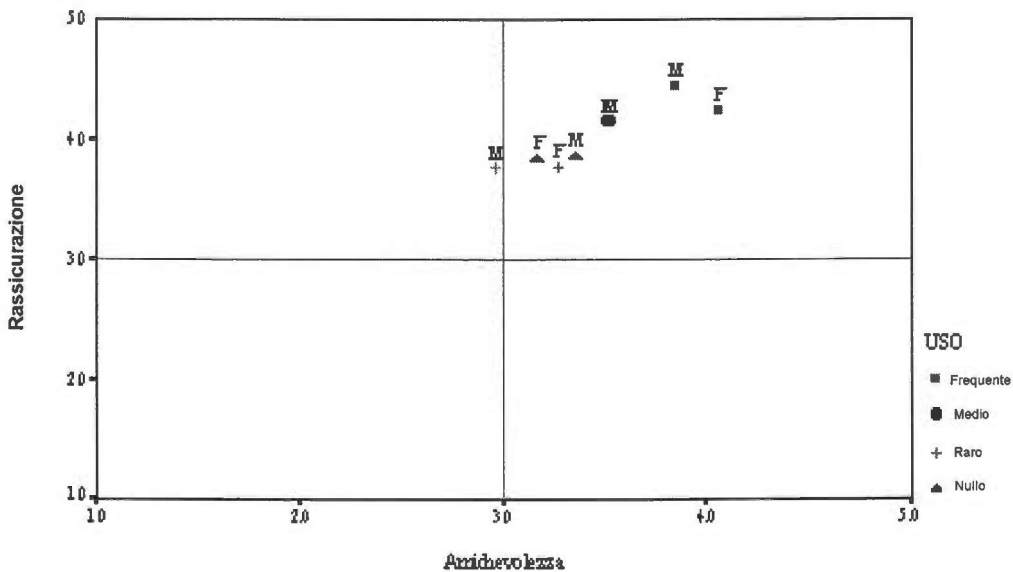


Figura 8. Giudizi di rassicurazione e amichevolezza in funzione di Uso e Sesso biologico.

Il trend comune dei due sessi e la correlazione fra i due fattori sono evidenti. Al crescere dell'esperienza aumentano entrambe le dimensioni. L'analisi post-hoc sul campione complessivo tramite il metodo di Scheffè evidenzia le seguenti differenze ($p < .05$): elevato vs. scarso e nullo, medio vs. scarso per quanto riguarda la rassicurazione; elevato vs. scarso e nullo per quanto riguarda la amichevolezza.

Un'ultima verifica ha riguardato il consenso attribuito allo stereotipo che definisce gli uomini più bravi delle donne nell'utilizzare il computer. Su tale asserzione fa perno la spiegazione disposizionalista. In generale, il campione si è dimostrato nettamente contrario (mediana = 5, assolutamente falso). Si evidenzia però una differenza significativa determinata dal Sesso Biologico dell'intervistato: gli uomini sono un po' più propensi ad accettare la veridicità dell'affermazione $U = 6585.5$, ($n = 259$), $p < .001$.

5.1.2.3. Predittori dell'esperienza

Per evidenziare le variabili che meglio permettono di prevedere l'esperienza informatica è stata eseguita un'analisi standard di regressione multipla, usando \sqrt{GDE} come variabile dipendente e amichevolezza, \log_{10} rassicurazione, Sesso Biologico e fascia di età (1, 18-19 anni; 2, 20-24; 3, oltre i 24) come predittori. Applicando la procedura della distanza di Mahalanobis, non sono stati individuati *outlier* multivariati, a un valore di probabilità pari a $p < .001$. La Tabella 5 riporta gli indici di correlazione fra le variabili, i coefficienti di correlazione non standardizzati B e i coefficienti di regressione standardizzati β . Il valore di correlazione multipla R è significativamente diverso da 0, $R = .45$, $F(3, 253) = 15.65$, $p < .001$. I valori di correlazione multipla quadrata R^2 sono rispettivamente .21 e .20 con correzione.

	GDE	Piace	Paura	B	β
Piace	.33			.11	.23***
Paura	.31	.34		-.30	-.21***
Sesso	-.21	-.05	.10	-.12	-.15**
Fascia	.18	.10	-.02	.07	.12*

Tabella 5. Regressione multipla di amichevolezza, rassicurazione e sesso su GDE.

Nelle prime tre colonne sono riportati i valori di correlazione fra le variabili. Legenda: *** $p < .001$ ** $p < .01$ * $p < .05$.

Tutte le variabili contribuiscono significativamente a predire l'esperienza informatica, anche se con forza diversa. Nel complesso la percentuale di varianza predetta, non è molto elevata, ma sembra chiaro il contributo di entrambi i fattori sottostanti l'atteggiamento verso il computer e del sesso. La percentuale di varianza spiegata aumenta notevolmente ($R = .40$) se si esegue la stessa analisi introducendo come predittori atteggiamento generale, sesso e indice di conoscenza, previa trasformazione di quest'ultimo in variabile dicotomica (1 nessuna, 2 minima).

5.1.3. Conclusioni

A conclusione di questo primo studio possiamo fare le seguenti considerazioni.

1. Il questionario EC appare uno strumento adeguato di autovalutazione dell'esperienza verso il computer. Si è infatti dimostrato efficiente nel raccogliere giudizi congruenti (solo il 3% del campione è stato considerato inaffidabile) e altamente correlati alle misure oggettive. Inoltre permette di discriminare chiaramente gli intervistati sulla base delle variabili evidenziate dalla letteratura.
2. Il questionario CATT offre buone prestazioni anche nella versione italiana, per quanto richiede una riduzione del set di item iniziale. Sembra inoltre che l'atteggiamento verso il computer si articoli su due dimensioni, un sentimento di *amichevolezza* e uno di *rassicurazione*.
3. L'ipotesi situazionalista appare più consona a spiegare i risultati ottenuti dal campione in esame. La minore presenza delle donne nel mondo informatico non è attribuibile a predisposizioni interne: l'atteggiamento verso il computer infatti è modulato in funzione dell'esperienza, ma non del sesso, nè dell'Identità Sessuale. Inoltre sembra plausibile l'ipotesi di un maggiore incoraggiamento sociale verso i maschi. La prospettiva situazionalista può inoltre essere estesa a spiegare l'atteggiamento più negativo degli studenti italiani rispetto ai loro colleghi americani precedentemente evidenziate in letteratura (Sensales e Greenfield, 1991).
4. In un campione caratterizzato da esperienza medio bassa atteggiamento e sesso possono essere considerati buoni predittori dell'esperienza. Anche l'età sembra contribuire, per quanto in maniera minore. La conoscenza comunque è il fattore in assoluto più predittivo. In generale i soggetti con maggiore esperienza sono quelli con atteggiamento più positivo, sono maschi e frequentano il biennio.

5.2. *Studio 2*

Il lavoro si propone essenzialmente di validare i risultati dello studio 1 su un campione caratterizzato da età ed esperienza diverse e di verificare nel dettaglio l'influenza dell'età sull'atteggiamento. A questo fine è stata testata una popolazione di studenti del primo e del penultimo anno di un Istituto Tecnico di Trieste e, come controllo, un campione di studenti dell'Università della Terza Età.

Quali ipotesi di lavoro è stata assunta l'esistenza dei seguenti effetti:

1. Particolare modulazione di esperienza e atteggiamento in funzione dell'età. Fra gli studenti di scuola media superiore ci si attende una differenza significativa sull'esperienza a favore dei più grandi, mentre si ipotizza che gli intervistati dell'Università della terza età abbiano in assoluto minore esperienza e atteggiamento più negativo.
2. Congruentemente con l'ipotesi situazionalista si assume che in un campione caratterizzato da maggiore esperienza informatica l'effetto del sesso sull'esperienza vada affievolendosi.

5.2.1. *Metodo*

I dati sono stati raccolti nell'autunno del 1993, utilizzando gli strumenti psicometrici precedentemente descritti, aggiornati però sulla scia delle modifiche suggerite dallo studio

1. La procedura seguita nel lavoro è identica alla precedente (§ 5.1.1.3).

5.2.1.1. *Popolazione della ricerca*

Il set di questionari è stato somministrato a un campione complessivo di 231 persone. Gli studenti di scuola media superiore sono 114 (72 F e 42 M), di cui 52 iscritti al primo anno e 62 al penultimo. Al momento della somministrazione, la loro età media era di circa 16 anni, da un minimo di 14 a un massimo di 20. Il campione degli studenti della terza età è invece costituito da 117 persone (78 F e 39 M), con età media di circa 67 anni, da un minimo di 38 a un massimo di 85.

5.2.2. Risultati

5.2.2.1. Verifica degli strumenti

Data la notevole differenza fra i due campioni oggetto di studio la verifica degli strumenti psicometrici è stata effettuata separatamente per ogni campione.

Questionario EC

Come era prevedibile, il campione della terza età è composto prevalentemente da persone che non hanno mai utilizzato un computer (77%). Questo rende difficile il lavoro di categorizzazione dell'esperienza informatica e impossibile la validazione dello strumento.

Ben diversa è la realtà degli studenti di scuola media superiore. Si evidenziano solo 4 non utenti (3 F, 1 M), che sono stati esclusi dal campione. Per quanto riguarda la consistenza interna della parte di autovalutazione, lo strumento ha nuovamente fornito prestazioni soddisfacenti, anche se lievemente inferiori a quelle evidenziate sul campione universitario.

Il confronto fra il giudizio diretto di esperienza (item 1; GDE) e la stima di tale valore ottenuta mediando le risposte agli item specifici (2-9; *giudizio indiretto esperienza*, GIE) evidenzia una correlazione significativa, $r = .72$, ($n = 104$), $p < .001$. Soddisfacente è anche la correlazione fra le risposte di esperienza e conoscenza della programmazione $r = .67$, ($n = 107$), $p < .001$. Lo screening del campione è stato eseguito seguendo le stesse procedure adottate nello studio 1, anche se, a causa della lieve tendenza a una minore congruenza nelle risposte e delle minori dimensioni del campione, la rigidità dei criteri di esclusione è stata lievemente allargata (esperienza-conoscenza programmazione > 2). In totale sono stati esclusi 6 intervistati, pari al 5% dell'intero campione.

La correlazione fra GDE e IOC (indice oggettivo di conoscenza, calcolato secondo la procedura descritta nello studio 1) risulta significativa, ma non del tutto soddisfacente $r = .53$, ($n = 77$), $p < .001$. Migliore invece è quella fra GIE e IOC $r = .68$, ($n = 74$), $p < .001$. Ciò induce a scegliere il giudizio indiretto come *indice generale di esperienza*. GIE è stato quindi normalizzato tramite calcolo della radice quadrata \sqrt{GIE} . L'indice d'*Uso* (nullo, scarso, medio e frequente) è stato calcolato come nello studio precedente, anche se sulla base di applicazioni diverse (item1, item3, item4, item6, item8 e item10).

Questionario CATT

Sul campione della terza età il CATT ha dimostrato un'elevata attendibilità ($\alpha = .83$), mentre nel caso degli studenti di scuola media superiore la coesione interna degli

item è risultata sensibilmente inferiore ($\alpha = .70$). Analizziamo per primo il caso più problematico.

Dall'analisi della correlazione item-totale si evidenziano in particolare due affermazioni (15 e 17)³⁶ che sembrano completamente estranei al resto della scala ($r < .10$). La loro esclusione produce una lieve crescita della coesione interna ($\alpha = .73$). La prestazione mediocre dello strumento può accreditarsi alla differenza d'età del particolare campione in analisi, inferiore a quelli su cui era stato originariamente valutato e poi tarato nella versione italiana. Infatti, eseguendo due separate analisi di attendibilità, una sui punteggi degli studenti del primo anno e l'altra su quelli del penultimo, si evidenzia una maggiore coesione interna al crescere dell'età (nell'ordine $\alpha = .68$ e $\alpha = .75$).

La struttura fattoriale dello strumento è stata analizzata tramite il metodo della probabilità massima (*maximum likelihood*) imponendo una configurazione bifattoriale, come suggerito dallo studio 1. Il modello sembra adeguato alle osservazioni $\chi^2(76) = 87.22, p = .18$. I due fattori spiegano complessivamente il 34.3% della varianza (il primo ha un autovalore di 3.4 ed è responsabile del 22.1% della varianza, il secondo ha un autovalore di 1.8). Gli item però si distribuiscono in modo diverso. Secondo un criterio che sancisce l'esclusione di tutti gli aggettivi spuri (entrambe le saturazioni $< .38$), e misti (una saturazione $> .30$ e l'altra $> .40$) il primo fattore risulta composto da 8 item (1, 2, 3, 6, 14, 16, 18, 19) mentre il secondo solo da 2 (8, 9). Ciò malgrado è ancora possibile evidenziare la divisione semantica precedentemente determinata. In generale il primo fattore (*amichevolezza*) fa riferimento alle caratteristiche positive dell'elaboratore elettronico, che viene descritto come attraente, utile e positivo. Il secondo è ancora chiaramente riferibile al senso di *rassicurazione* indotto dal computer. Si noti infatti che i due item da cui è composto corrispondono a quelli che saturavano massimamente il primo fattore dello studio 1, mentre i tre (6, 18, 19) che si sono spostati al fattore della amichevolezza sono effettivamente ambigui e possono dare luogo a entrambe le interpretazioni.

Analizziamo adesso il campione della terza età. L'elevato indice di attendibilità del CATT viene ulteriormente accresciuto ($\alpha = .83$) eliminando 2 item (5 e 17) che appaiono relativamente poco correlati al resto della scala ($r < .20$). La configurazione bifattoriale non è del tutto adeguata ai dati, $\chi^2(76) = 112.27, p < .05$, eppure il paragone con il valore di χ^2 dell'ipotesi unifattoriale e l'analisi dello *screen-test* suggeriscono che sia tuttora la soluzione migliore. I due fattori spiegano complessivamente il 42.7% della varianza. Il

³⁶Per permettere il paragone con i risultati dello studio 1 gli item sono etichettati con gli stessi numeri.

primo corrisponde alla *rassicurazione* e satura gli item 3, 4, 8 e 9; il secondo alla *amichevolezza* (item 1, 2, 7, 14). E' evidente che il campione della terza età riproduce più da vicino i risultati ottenuti dagli studenti universitari.

L'analisi dell'attendibilità sul campione completo ($\alpha = .79$) suggerisce di escludere l'item17 ($\alpha = .80$).

Questionario BSRI

La classificazione dei soggetti in *femminili*, *maschili*, *androgini* e *indifferenziati* è stata eseguita sulla base delle procedure utilizzate nello studio 1. I due campioni sono stati distribuiti sulla base della posizione dei punteggi F e M, rispetto alle mediane del loro specifico gruppo. La distribuzione dei due campioni è riportata in Tabella 6.

Identità Sessuale	Scuola Media Superiore	Università della terza età
Androgini	21%	19%
Indifferenziati	23%	21%
Maschili	29%	30%
Femminili	27%	30%

Tabella 6. Distribuzione delle categorie di Identità Sessuale nei due campioni.

5.2.2.2. Analisi del campione di scuola media inferiore

Abbiamo già visto che il campione della terza età è prevalentemente costituito da persone che non hanno mai utilizzato un computer. Di conseguenza la parte dell'analisi sulla modulazione dell'esperienza può riferirsi esclusivamente agli studenti di scuola media superiore.

Come previsto, i risultati relativi all'alfabetizzazione informatica sono migliori di quelli del campione di universitari. Il 21% degli intervistati dichiara di utilizzare il computer *raramente*, il 31% *a volte*, il 31% *frequentemente* e il 17% *molto spesso*. È interessante inoltre notare che solo il 20% ammette di non saperlo usare autonomamente. Principale responsabile della migliore conoscenza informatica è sicuramente la crescente politica di informatizzazione che sta caratterizzando la scuola. Il 21% degli intervistati ha infatti dichiarato di avere seguito dei corsi di informatica già alle medie inferiori, e lo 88% alle superiori (contro rispettivamente il 6% e il 21% del campione universitario). Per

evitare generalizzazioni inadeguate, si ricordi che gli studenti erano stati campionati da un Istituto Tecnico.

In generale gli intervistati sono concordi nell'affermare di volere aumentare la propria conoscenza informatica (media = 3.91, SE = .09). Nuovamente, in contrapposizione con l'ipotesi disposizionalista, non si evidenzia un effetto del Sesso Biologico. Sia per le femmine che per i maschi il desiderio di migliorare la propria conoscenza aumenta significativamente al crescere dell'esperienza [$r = .26$, ($n = 90$), $p < .05$].

Il grafico in Figura 9 riporta i valori medi di utilizzo delle otto applicazioni previste nel questionario. Nuovamente l'applicazione principale del computer consiste nel gioco, poi, in ordine di frequenza, nella programmazione, videoscrittura, calcolo e disegno. La bassa frequenza d'uso delle due applicazioni tipicamente legata alle reti (posta elettronica e IR) rispecchia l'arretratezza tecnologica della società italiana rispetto a quella americana. Come previsto dall'ipotesi situazionalista, secondo i risultati di un'analisi di Mann-Whitney la differenza fra maschi e femmine emerge in modo deciso nel gioco $U = 881.5$ ($n = 102$), $p < .01$, mentre tende a scomparire nell'uso delle altre applicazioni. L'unica eccezione riguarda l'utilizzo dei fogli elettronici tendenzialmente superiore nel campione dei maschi $U = 981$ ($n = 100$), $p = .07$.

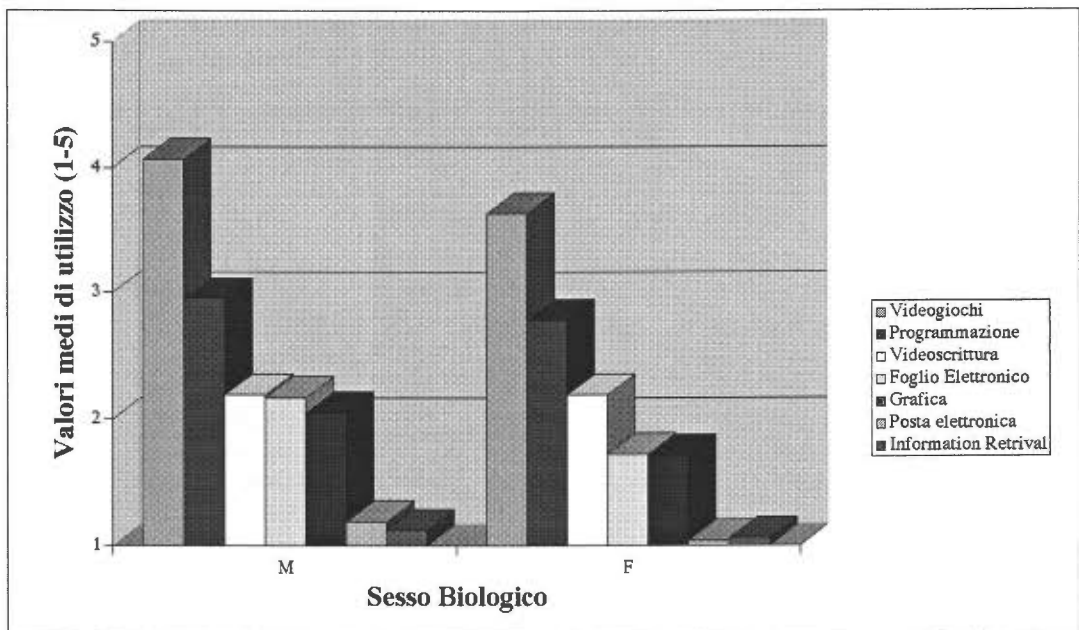


Figura 9. Utilizzo di applicazioni informatiche in funzione del Sesso Biologico.

Possiamo dunque concludere che il sesso non sembra avere grande influenza sulla frequenza d'uso delle singole applicazioni. L'effetto dell'Età invece emerge chiaramente ($p < .05$) su tutte le applicazioni tranne computer grafica, videogiochi e IR (vedi Figura 10).

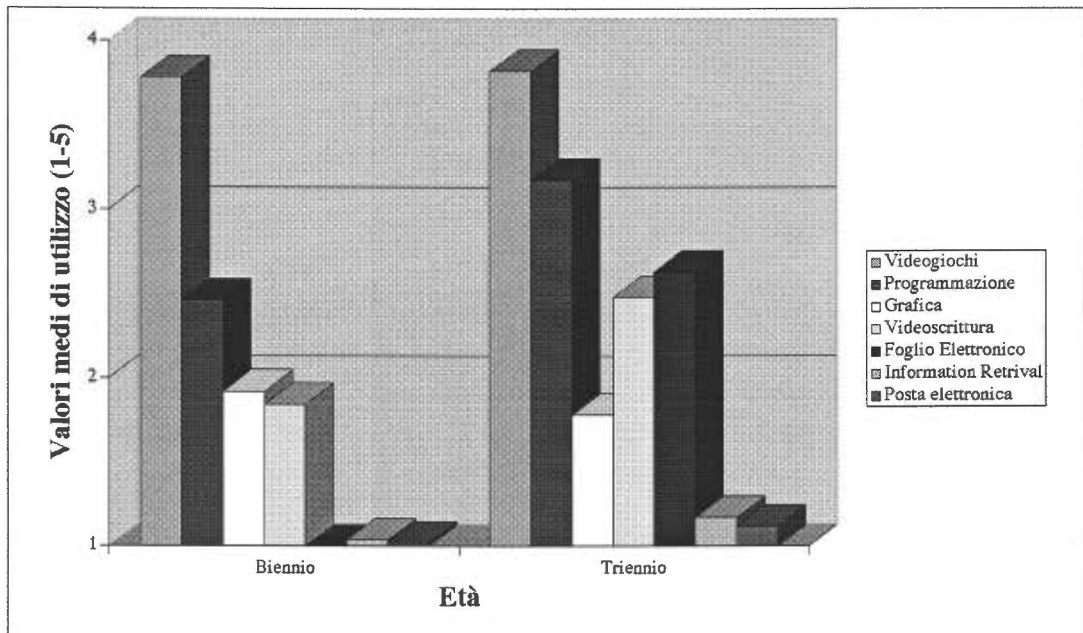


Figura 10. Utilizzo di applicazioni informatiche in funzione dell'Età.

In media l'atteggiamento generale degli intervistati è risultato abbastanza positivo (media = 3.90 SE = .07), così come il giudizio sulla amichevolezza (media = 3.93, SE = .06). Relativamente al fattore rassicurazione l'opinione generale è chiara: il 75% del campione dichiara di non avere la minima paura del computer. Ai fini delle analisi statistiche dunque viene utilizzato esclusivamente l'indice di atteggiamento generale, che oltre a comprendere entrambe le dimensioni ha anche il vantaggio di distribuirsi più normalmente.

Atteggiamento e \sqrt{GIE} sono state analizzate mediante una MANOVA, per il disegno Età (2) * Sesso (2). A livello del confronto multivariato emerge un marcato effetto dell'Età, $F(2, 85) = 7.21, p = .001$. Come si può chiaramente vedere dalla Figura 11 tale effetto è causato da una notevole crescita nell'esperienza, $F(1, 86) = 13.8, p < .001$, mentre l'atteggiamento in sé non subisce modifiche, $F(1, 86) < 1$. L'effetto principale del Sesso non è significativo $F(1, 86) < 1$. A livello univariato emerge una tendenziale interazione Età * Sesso, $F(1,86) = 3.59, p = .06$, che può essere attribuita a un lieve calo di esperienza delle femmine più grandi rispetto ai loro coetanei. Questi risultati rappresentano dunque un'ennesima conferma dell'ipotesi situazionalista.

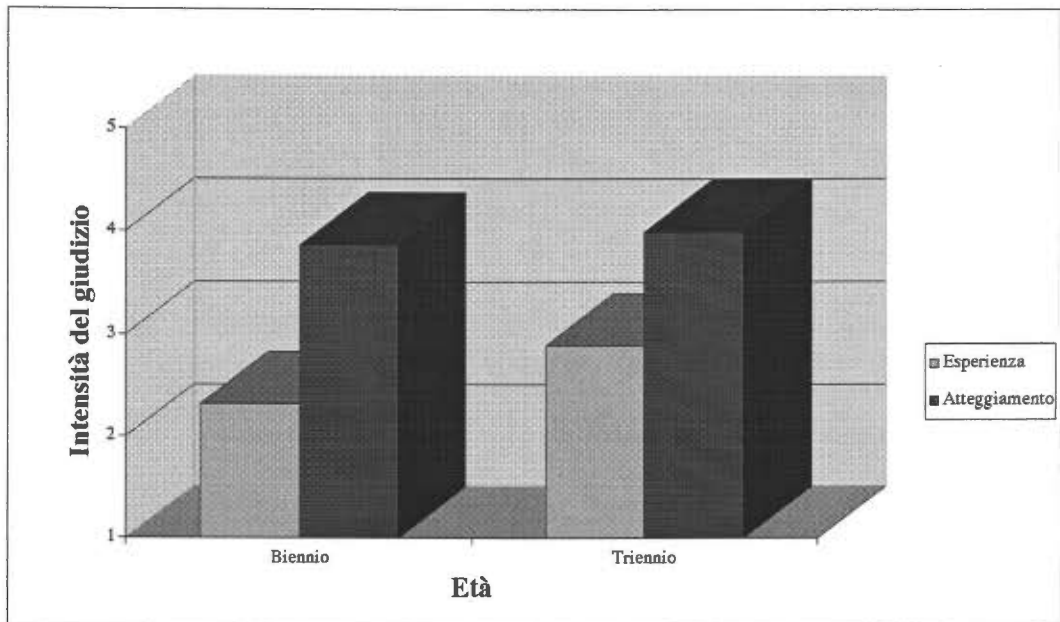


Figura 11. GDE e Atteggiamento in funzione dell'Età.

La stessa analisi sul disegno Età (2) * Identità Sessuale (4) mostra un tendenziale effetto dell'identità sessuale sull'atteggiamento $F(3, 80) = 2.51, p = .06$. L'analisi post-hoc, mediante il metodo di Duncan ($p < .05$), evidenzia che la differenza è dovuta al relativo calo di atteggiamento nei partecipanti indifferenziati, rispetto a quelli maschili e femminili.

5.2.2.3. Predittori dell'esperienza

Per evidenziare le variabili che meglio permettono di prevedere l'esperienza è stata eseguita un'analisi standard di regressione multipla, con \sqrt{GIE} come variabile dipendente e Atteggiamento, Sesso e Età come predittori. I principali risultati sono riportati nella Tabella 7. Il valore di correlazione multipla R è significativamente diverso da 0, $R = .58$, $F(3, 86) = 14.52, p < .001$.

È interessante vedere che in questo particolare campione, caratterizzato da esperienza medio alta, il sesso biologico non predice l'esperienza. Infatti le uniche variabili che determinano largamente il valore dell'esperienza sono Atteggiamento e Età. Il valore di correlazione multipla quadrata è soddisfacente ($R^2 .34$), testimoniando che le due variabili hanno insieme una forte capacità predittiva, per quanto la principale è sicuramente l'atteggiamento.

	GDE	Atteggiamento	B	β
Atteggiamento	.48		.16	.42 ***
Sesso	-.23	-.17	-.04	-.09
Età	.36	.13	.12	-.28**

Tabella 7. Risultati analisi regressione

Le prime due colonne riportano gli indici di correlazione, la seconda i coefficienti di correlazione non standardizzati B e i coefficienti di regressione standardizzati β . Legenda: *** $p < .001$ ** $p < .01$.

5.2.2.4. Paragone fra i due campioni

Per permettere il paragone fra i due campioni, è stato calcolato un indice di atteggiamento comune, sulla base dei risultati dell'analisi di attendibilità eseguita sul campione completo. Per rispettare il requisito dell'omogeneità di varianza tale indice è stato elevato al cubo e poi analizzato mediante un'ANOVA a 2 fattori, Campione (2) * Sesso Biologico (2). Come previsto, l'effetto principale del campione è estremamente significativo $F(1, 177) = 18.47, p < .001$. Emerge anche un chiaro effetto del Sesso $F(1, 177) = 7.27, p < .01$, a favore dei maschi. Non risulta significativa invece l'interazione [$F(1,177) < 1$]. Gli effetti semplici permettono di evidenziare che l'effetto del sesso è interamente dovuto al campione della terza età $F(1,83) = 3.99, p < .05$. Questo risultato parzialmente contraddice l'ipotesi situazionalista.

5.2.3. Conclusioni

Dalla ricerca emergono i seguenti risultati.

- Si conferma la validità interna, esterna e di costrutto del questionario EC.
- Il CATT non appare particolarmente efficiente se applicato a una popolazione di studenti giovani, mentre al crescere dell'età il suo valore viene confermato. In generale l'ipotesi bifattoriale sembra l'interpretazione migliore della struttura interna del costrutto misurato dal reattivo.

- Nuovamente l'ipotesi situazionalista sembra la spiegazione migliore del quadro di risultati ottenuto. Nel campione degli studenti di scuola superiore il gap fra uomini e donne, non solo non è attribuibile a predisposizioni interne, ma tende a scomparire. Se l'ambiente circostante non è discriminatorio, le ragazze mostrano un comportamento simile a quello dei coetanei. L'unico dato a favore dell'ipotesi disposizionalista è l'effetto significativo del sesso sull'atteggiamento nel campione della terza età. Questo risultato può essere facilmente spiegato se si considera che le persone più anziane sono anche quelle che maggiormente adeguano il proprio comportamento e la propria visione del mondo agli stereotipi sessuali.
- Si conferma la modulazione dell'effetto dell'Età sull'esperienza come previsto dalla ipotesi di lavoro. Se il contesto esterno è positivo, vale a dire interessato all'aumento dell'alfabetizzazione informatica, al crescere dell'età aumenta l'esperienza. Gli studenti del penultimo anno mostrano infatti più esperienza di quelli del primo, ma anche più esperienza del campione universitario. L'approccio situazionalista sembra dunque più adeguato anche a spiegare l'effetto dell'Età. In questa visione, il problema della analfabetizzazione informatica della terza età viene quindi attribuito oltre che al supposto decrescere delle abilità cognitive anche alla mancanza di stimoli esterni.
- In un campione caratterizzato da esperienza medio alta atteggiamento e età si dimostrano buoni predittori dell'esperienza. Il sesso invece non è adeguato allo scopo.

5.3. *Discussione*

In sintesi le due ricerche vanno a favore dell'ipotesi situazionalista sottostante l'approccio globale, che postula una modulazione dell'effetto delle variabili anagrafiche sulla base dell'esperienza e dell'atteggiamento delle persone.

L'importanza pratica di questo lavoro appare evidente considerando le diverse conseguenze che l'ipotesi disposizionalista e quella situazionalista possono avere sulla progettazione e sulla valutazione di sistemi informatici. Se si accetta la prima sesso e età devono venire considerati determinanti fondamentali e non ambigue del comportamento di interazione. Se si accetta la seconda, invece, entrambe le variabili appaiono meno importanti, in quanto il loro effetto è mediato da fattori esterni.

Il lavoro conferma dunque la necessità di raccogliere il maggiore numero di informazioni possibili sull'utente. A questo fine vengono forniti due strumenti che dimostrano un'elevata validità. Il questionario EC può facilmente essere abbreviato, togliendo la parte di valutazione oggettiva della conoscenza che ha sempre dimostrato elevati valori di correlazione con l'indice di esperienza autovalutata tramite stima della frequenza d'uso. In questo modo si ottiene uno strumento agile e veloce adatto a discriminare gli utenti prima dell'effettiva valutazione.

6. Simulazione 1

I due esperimenti di simulazione presentati in questa tesi si occupano dell'analisi di sistemi per la compilazione di moduli, in cui l'interazione si basa su linguaggio naturale, mediato da tastiera, e su una forma elementare di gestualità: il puntamento, mediato dal mouse.

La configurazione dei sistemi analizzati è estremamente interessante poiché riproduce una combinazione *artificiale* di modalità (Schmauks, 1987; Cohen e Sullivan, 1989; Schmauks e Willie, 1991; Stock, Strapparava e Zancanaro, 1995; De Angeli *et al.*, 1996). Indicare un oggetto mentre si scrive è estremamente diverso da indicare un oggetto mentre si parla. Analizziamo in breve la situazione naturale.

Nell'interazione fra umani l'identificazione di referenti nello spazio può essere effettuata in vari modi che comprendono nominazione diretta, metonimia, riferimenti anaforici o descrizione dell'oggetto target. Un metodo frequente ed efficace è la deissi spaziale³⁷ con cui il parlante localizza il target ponendosi come centro del sistema di riferimento e supportando il linguaggio, nella veste di un apposito termine deittico (questo e quello; qui e lì; qua e là) con un *gesto*, quale l'estensione del dito indice della mano dominante nella direzione del target. La deissi permette di descrivere la complessità dell'esperienza percettiva meglio del solo linguaggio, specialmente nei casi in cui la nominazione diretta non sia possibile (McNeill, 1992; Glenberg e McDaniel, 1992; Gullberg, 1996). Tale strategia rappresenta dunque un interessante esempio di distribuzione di caratteristiche semantiche fra gestualità e linguaggio (McNeill, 1992; Cassel e Prevost, 1996). Il significato trasmesso dalle due modalità non è ridondante ma ciascuna contribuisce in modo unico e necessario alla costituzione del significato finale. La procedura naturale di attuazione prevede la produzione parallela e altamente sincronizzata di termini e gesti deittici (Levelt, Richardson e La Heij, 1985). Tale processo solitamente è determinato in modo inconsapevole (Kendon, 1980). Secondo la teoria balistica (Levelt *et al.*, 1985), le due modalità interagiscono a livello di pianificazione, ma sono poi caratterizzate da un processo di esecuzione modulare. In caso di modifica esterna della situazione è il parlato che tende ad adattarsi al gesto.

³⁷ si noti che in sebbene le deissi possano avere anche un uso simbolico in questa tesi ci si riferisce esclusivamente al caso in cui tale strategia di identificazione è accompagnata dalla gestualità (Levinson, 1983).

Analizziamo adesso la situazione artificiale. I sistemi multimodali, come quelli oggetto di studio, implicano la sostituzione della procedura parallela con una di tipo sequenziale. La modalità ostensiva e quella verbale sono mediate dallo stesso effettore, la mano che deve spostarsi fra tastiera e mouse. Inoltre, per semplificare l'attribuzione del gesto alla sua controparte linguistica gli attuali prototipi richiedono di effettuare il puntamento secondo una procedura estremamente vincolata, puntando subito prima o subito dopo la scrittura del termine deittico. La modifica procedurale potrebbe così annullare il vantaggio della comunicazione multimodale e rendere i sistemi in questione inutili. Diventa dunque necessario analizzare empiricamente le problematiche legate al loro utilizzo.

Le due simulazioni sono finalizzate a dimostrare da punti di vista diversi la superiorità della comunicazione multimodale su quella unimodale malgrado i vincoli imposti dal partner artificiale. L'*ipotesi del vantaggio* si basa essenzialmente sulla maggiore flessibilità dialogica per cui i sistemi multimodali dovrebbero essere migliori e soddisfare maggiormente gli utenti. In questo gli studi sono congruenti con l'indirizzo di ricerca prevalente del settore (Oviatt e Olsen, 1994; Oviatt, et al., 1995; Oviatt, 1996; Oviatt *et. al.*, 1997).

Entrambi i lavori però affrontano l'argomento secondo l'ottica dell'approccio globale. Innanzitutto si pongono in contrasto con l'ipotesi ingenua per cui l'introduzione di una forma di comunicazione umanoide dovrebbe permettere un immediato e indifferenziato accesso alla tecnologia. Si ritiene, infatti, che malgrado la maggiore flessibilità anche nell'interazione multimodale il computer continui a rimanere un partner '*speciale*'. L'interazione è diversa dalla comunicazione fra umani e quindi lo scambio richiede del tempo per raggiungere livelli ottimali di interazione (*ipotesi dell'adattamento*). Questo periodo è variabile a seconde di specifiche caratteristiche dell'utente, alcune delle quali possono essere previste (*ipotesi delle differenze individuali*).

In breve, i lavori si occupano di evidenziare i fattori che agiscono sull'interazione sia quando è a regime che nella prima fase di adattamento. In questa prospettiva si cerca anche di chiarire le relazioni strutturali fra misure oggettive e soggettive della prestazione, secondo l'idea che l'utente non solo sia un buon giudice, ma rappresenti anche arbitro finale del successo o dell'insuccesso del sistema (*ipotesi della congruenza*).

Descriviamo adesso nel dettaglio la prima simulazione sulla base dello schema proposto dall'approccio globale (§ 4.1.1)³⁸.

³⁸ Un sottoinsieme dei risultati è stato presentato nello scorso ottobre al *Workshop on the Integration of Gesture in Language and Speech* svoltosi a Newark e Wilmington Delaware (USA). In quel contesto ci si era concentrati unicamente sull'analisi dei comportamenti comunicativi sottostanti l'identificazione di referenti nello spazio; in questo capitolo, invece, viene proposta una dettagliata analisi di usabilità.

6.1. Definizione degli obiettivi

Lo studio costituisce la valutazione comparata tramite simulazione semiautomatica di alcune soluzioni di interfaccia multimodale. Il prototipo oggetto di studio, identificato dall'acronimo MIS (*Multimodal Intelligent System*), capace di aiutare gli utenti nella compilazione di moduli. In particolare sono stati analizzati due aspetti dell'interfaccia: uno relativo all'apparenza del modulo, l'altro relativo al feedback associato all'azione di puntamento.

- *Apparenza del modulo.*

I moduli si differenziano sulla base dell'Identificabilità dei Campi (IC). Nella condizione IC *Completa*, tutti i campi sono identificati da una diversa lettera, facilitando così il ricorso alla nominazione diretta (per esempio: "campo A"). Nella condizione IC *Parziale* solo un campo su tre è etichettato, mentre gli altri sono identificati da asterischi (vedi Figura 12). In quest'ultimo caso l'identificazione linguistica richiede di specificare la posizione del target sulla base dei campi etichettati (per esempio: "il campo a destra di A").

- *Feedback.*

Nella condizione *Feedback Presente* il puntamento è immediatamente seguito dal cambiamento di colore del campo selezionato (vedi Figura 12), mentre nella condizione *Feedback Assente* non si verifica alcun cambiamento. Apparentemente le due soluzioni si basano su strategie di progettazione opposte. La prima è congruente con una delle direttive generali proposte da Shneiderman (1992; § 3.2) per cui a ogni azione dell'utente deve sempre seguire un chiaro feedback del sistema. L'altra, invece, asserisce che l'interazione utente-computer dovrebbe ricalcare da vicino la comunicazione fra umani (Schmauks, 1987).

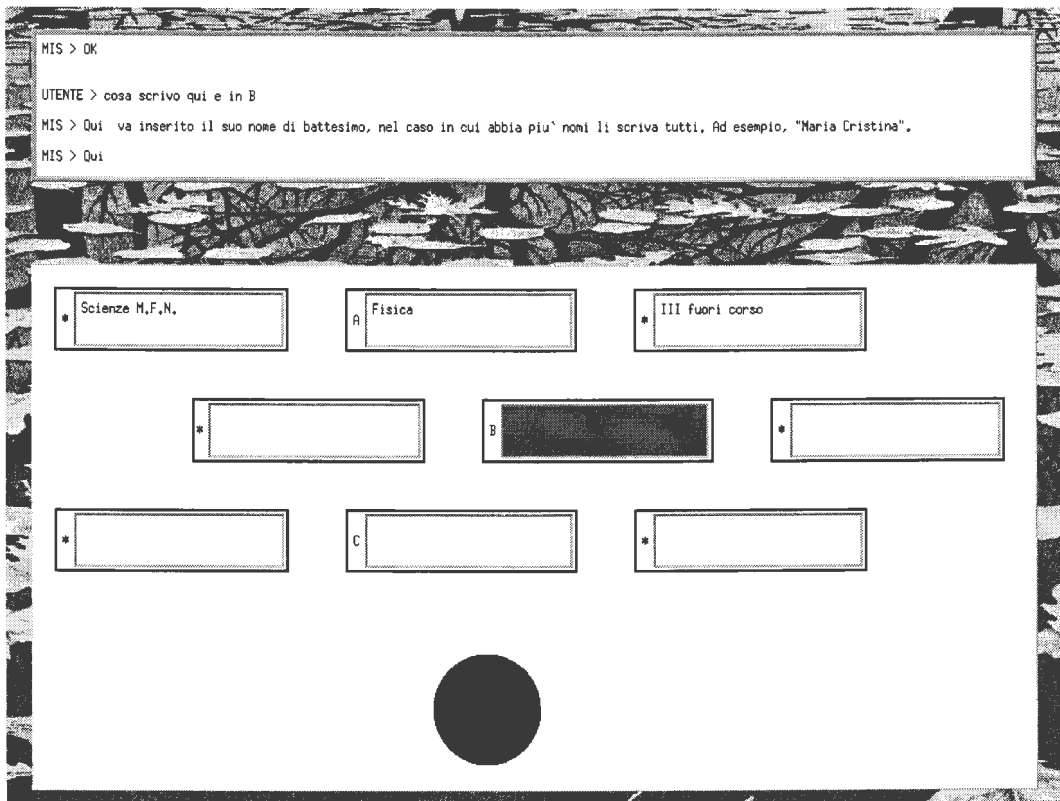


Figura 12. Interfaccia dell'utente nella condizione IC Parziale e Feedback Presente.

Le ipotesi di lavoro si articolano su vari livelli di specificità. Consideriamo innanzitutto le ipotesi più generali precedentemente descritte. Quelle relative all'usabilità verranno descritte più avanti in corrispondenza all'operazionalizzazione degli attributi analizzati nello studio (§ 6.3).

Secondo l'*ipotesi del vantaggio* la preferenza naturale per la deissi non dovrebbe essere annullata dalla procedura cui gli utenti sono costretti. Ci si aspetta dunque che nella condizione IC Completa la deissi abbia la stessa probabilità di verificarsi della nominazione diretta, mentre diventi la strategia di comunicazione privilegiata nella condizione IC Parziale. A questo livello non si ipotizza alcun effetto del fattore Feedback che, invece, è considerato importante nel determinare il modo in cui l'espressione deittica viene prodotta.

A livello di produzione ci si aspetta che gli utenti cerchino di compensare la procedura non naturale agendo sulla formulazione dell'espressione deittica (*ipotesi della produzione*). Si ipotizza dunque che le espressioni vengano formulate in modo da ridurre al minimo i passaggi fra le diverse modalità. Il puntamento dovrebbe dunque essere effettuato di preferenza alla fine della frase subito dopo la scrittura del termine deittico. A questo livello viene ipotizzato anche un vantaggio relativo della condizione Feedback Presente, dove il cambiamento di colore potrebbe stimolare la produzione del puntamento all'interno della frase.

6.2. *Caratterizzazione delle quattro entità.*

Dal momento che il lavoro è orientato alla manipolazione di specifiche caratteristiche dell'artefatto sociale, particolare attenzione è stata prestata a mantenere costante l'effetto delle determinanti relative alle altre entità.

- *Ambiente.* La raccolta dati è stata eseguita in un laboratorio di usabilità in condizioni di illuminazione artificiale. Possiamo quindi assumere come minima e costante l'influenza del *contesto fisico*. L'interazione avveniva singolarmente e i partecipanti erano lasciati soli con il computer per tutta la durata della prova. Inoltre, le istruzioni specificavano chiaramente che gli sperimentatori non avevano alcun rapporto con l'implementazione del sistema. Il ruolo del *contesto sociale* può dunque venire considerato ridotto ai minimi termini e per quanto possibile costante.
- *Compito.* La scelta dello specifico compito, compilazione di moduli con informazioni socio-anagrafiche, è maturata sulla base di approfondite riflessioni. Innanzitutto è ecologico, in quanto rappresenta una delle applicazioni multimodali maggiormente analizzate negli ultimi anni (Schmauks, 1987; Cohen e Sullivan, 1989; Schmauks e Willie, 1991). Poi, richiede un'interfaccia altamente strutturata che permette di prevedere il comportamento dell'utente e di limitarne la variabilità linguistica (§ 6.2.1). La suddivisione del lavoro è la seguente: "MIS" fornisce le informazioni di compilazione, verifica in modo molto generale l'adeguatezza dei dati forniti dall'utente e procede all'inserimento dati. L'utente chiede le informazioni di compilazione, fornisce i dati richiesti e ne verifica la correttezza. A livello generale si può sostenere che la principale incongruenza fra *modello dell'utente* e *modello del progettista* riguarda un particolare aspetto del *compito informatico*. L'utente non può inserire direttamente i dati nel modulo, ma deve comunicarli a MIS che procede all'inserimento automatico. Questa scelta deriva dalla volontà di massimizzare l'uso del linguaggio naturale e di vincolare il mouse solo all'azione di puntamento, togliendogli il controllo del cursore. In generale comunque le due rappresentazioni appaiono abbastanza simili: come in una situazione naturale simile, ogni entità possiede delle conoscenze e ha bisogno di altre. Si può dunque prevedere che il compito sia semplice.

6.2.1. Caratterizzazione del computer

La simulazione girava su due SUN SPARC workstation collegate in rete mediante un software realizzato in MOTIF. Il tempo di trasmissione dell'informazione era ridotto al minimo: il mago monitorava il comportamento dell'utente *on-line* e spediva le sue risposte scegliendo fra un set di frasi pre-registrate.

L'interfaccia dell'utente è costituita da due finestre: una dedicata al *dialogo*, l'altra alla rappresentazione del *modulo* (vedi Figura 12). La tastiera è attiva solo nella prima in concomitanza ai turni dialogici dell'utente; il mouse è attivo solo nella finestra del modulo allo scopo di indicare i campi.

L'interfaccia del mago è invece più complessa (vedi Figura 13).

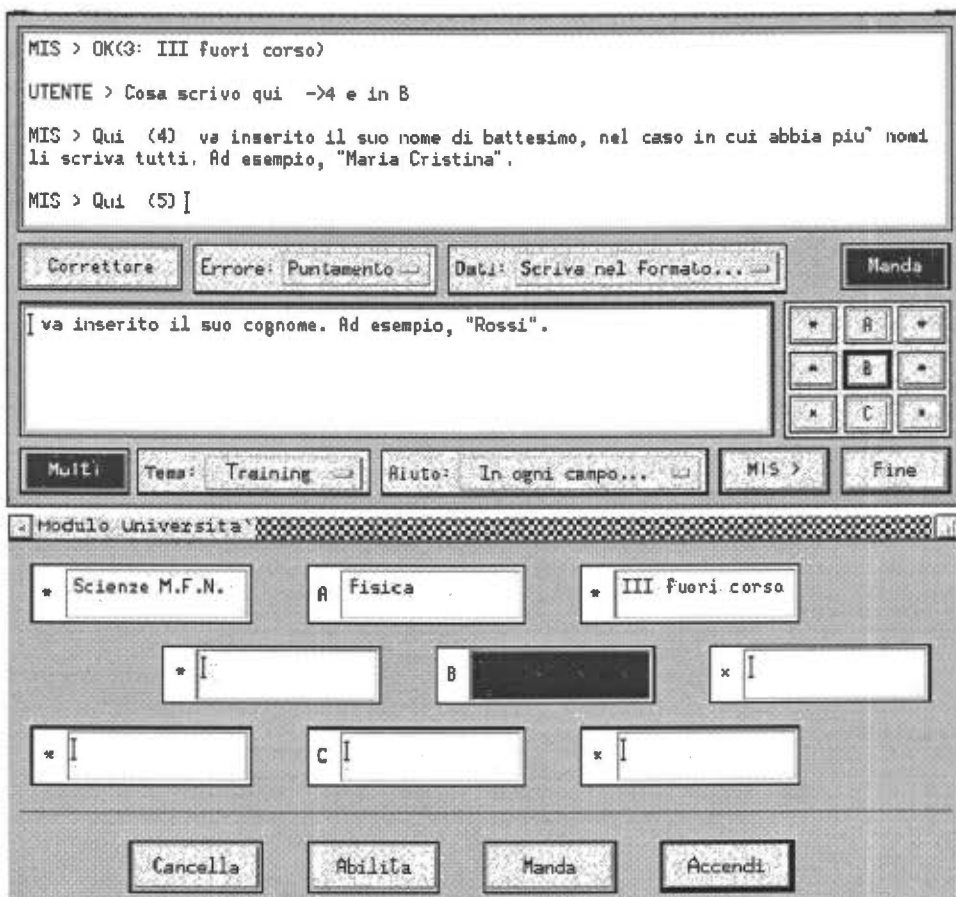


Figura 13. Interfaccia del mago nelle condizioni IC Parziali.

Si noti che i campi del modulo sono sfalsati rispetto all'ordinamento verticale: questo serve a impedire i riferimenti sulla dimensione verticale e inibire la percezione del modulo come un'entità composta di campi ordinati sequenzialmente dal primo al nono.

La finestra superiore fornisce una eco della finestra di dialogo dell'utente da cui il mago può monitorarne direttamente le azioni linguistiche. Quella immediatamente sotto

serve invece per generare le risposte, tipicamente prodotte in modo automatico tramite selezione di uno dei *tasti* da cui è circondata. L'ultima finestra in basso riproduce una eco del modulo dell'utente. Le informazioni vengono inserite mediante una sequenza taglia-incolla supportata dal mouse. Indipendentemente dalla condizione di Feedback il puntamento è sempre evidenziato dal cambiamento di colore del campo selezionato dall'utente oltre che dal riferimento numerico (-> 4) visibile sulla finestra di dialogo. I tasti inferiori servono alla gestione del puntamento. L'interfaccia è altamente sicura e robusta all'errore.

6.2.1.1. Modello interattivo

L'interazione riproduce un dialogo di tipo domanda-risposta, completamente diretto dall'utente. MIS infatti è poco cooperativo e direttivo. L'alternanza dei turni dialogici è rigidamente determinata entro uno schema di avvicendamento del tipo U-M-U-M (Sacks, Schegloff e Jefferson, 1974). Tale successione prevede che l'Utente scriva il suo messaggio e poi, pigiando il tasto ENTER, ceda il turno a MIS, il quale fa il suo intervento e riconsegna il turno all'Utente. MIS veniva simulato in modo da non saltare mai il turno. Dunque, la simmetria poteva essere interrotta solamente da un silenzio dell'utente. Queste situazioni, per altro assai rare, venivano risolte forzando il turno conversazionale e stimolando l'utente mediante una generica richiesta di azione. Le sovrapposizioni erano impossibili, poiché la tastiera dell'utente era abilitata solo durante il suo turno interattivo, chiaramente caratterizzato dalla presenza del prompt *UTENTE>* e dal lampeggiare del cursore sulla linea di dialogo (vedi Figura 12). Un sistema di filtri automatici impediva che il mago per errore inviasse messaggi fuori turno.

Il modello del comportamento interattivo di MIS è stato elaborato sulla base delle capacità degli attuali sistemi di comprensione del linguaggio naturale. In seguito alla prima formulazione (§ 4.2.3) e a studio pilota ($n = 6$) con il sistema di simulazione completo, i possibili interventi di MIS sono stati classificati in quattro categorie, a seconda del loro scopo all'interno del dialogo: (a) segnalazione di problemi comunicativi; (b) informazione sul contenuto dei campi; (c) informazione sulle proprie capacità o sul compito; (d) iniziative. Per ognuna di esse sono state quindi elaborate e registrate nell'interfaccia del mago alcune risposte standard.

Analizziamo in breve il primo punto. A livello lessicale MIS riconosce tutte le parole italiane scritte correttamente e correlate al dominio del compito. In caso di errore ortografico il mago attiva direttamente una finestra con la seguente frase "*Mi dispiace, non conosco la parola [parola errata]*", senza attendere che il partecipante completi il suo turno linguistico. A livello sintattico MIS riconosce solo frasi complete (le ellissi venivano

dunque respinte), grammaticalmente corrette (congruenza fra tutti gli elementi sintattici della frase) e non complesse (un solo riferimento anaforico e una sola subordinata per frase). In caso di mancato riconoscimento la risposta standard è *"Mi dispiace, non capisco la sua frase"*. A livello multimodale il riconoscimento è vincolato alle capacità dei prototipi multimodali, se il puntamento non compare in stretta successione al termine deittico l'errore viene segnalato con la frase *"Mi dispiace, non capisco dove vuole scrivere"*.

Tutte le *informazioni sul contenuto dei campi* erano state preregistrate. Il mago quindi deve solo selezionare il tasto corrispondente al campo desiderato. In questo modo spedisce una frase standard del tipo: *"Qui [selezione] vanno inserite le iniziali del suo nome e cognome, senza punto. Ad esempio, se lei si chiama Marco Rossi scriva 'MR'"*. MIS usa sempre deissi, evidenziate sull'interfaccia dell'utente dal cambiamento di colore del campo indicato. Quando il mago inserisce i dati anagrafici nel modulo, automaticamente viene generata nella finestra di dialogo la scritta *"MIS> OK"* e il turno passa all'utente. Se l'informazione fornita dall'utente è chiaramente implausibile il mago non la accetta e invia un messaggio del tipo: *"Dato non conforme alle mie conoscenze. Controlli la sua risposta"*. Questo messaggio viene inviato anche a seguito di una seconda risposta implausibile, mentre alla terza segue la frase *"Dato da verificare in seguito"* e l'informazione non viene inserita nel modulo. Se il partecipante fornisce un dato in modo non conforme ai canoni richiesti, il mago può accettarlo solo se lo scostamento è minimo altrimenti invia un messaggio di richiamo tipo: *"Scriva le informazioni nel formato richiesto"*.

Dato il particolare modello dialogico riprodotto le *informazioni sulle capacità o sul compito* e le *iniziative* di MIS erano estremamente limitate. Le prime venivano fornite solo in seguito a specifiche richieste dell'utente con una frase standard *"In ogni campo va inserita una diversa informazione. Se lei mi indica un campo, sono in grado di specificare quali dati vanno inseriti"* modificabile a seconda del contesto. Nei casi in cui un utente comunicava il suo disagio, la sua incapacità di eseguire il compito oppure se non scriveva niente per almeno cinque minuti il mago forzava il dialogo, prendeva il turno conversazionale con la frase *"Anche se non è sicuro, mi faccia comunque una domanda. Cercherò di aiutarla"*.

6.2.2. Caratterizzazione dell'utente

Data la generalità del sistema oggetto di studio, il campione è stato selezionato da una popolazione di studenti universitari pareggiando il numero di partecipanti sulla base del sesso. Le determinanti dell'utente sono state considerate con estrema cautela.

Consideriamo innanzitutto l'*agente cognitivo puro*. A livello dei processi primari di elaborazione, il compito prototipo richiede un carico cognitivo limitato (recupero di

informazioni familiari) e preventivamente bilanciato fra i compiti. A livello dei processi superiori l'interazione appare essenzialmente influenzata da caratteristiche dell'*agente socio-cognitivo* e dell'*agente individuale*, che sono state controllate.

Per controllare la *motivazione* i partecipanti sono stati pagati e la valutazione è stata presentata come l'analisi del sistema e non di prestazioni o abilità personali. Inoltre tutti i partecipanti erano sollecitati a completare tutti i compiti nel minor tempo possibile, in modo da evitare che alcuni si dedicassero più a scoprire le capacità dialogiche del sistema che a eseguire il compito. L'*esperienza informatica*, operazionalizzata come familiarità con i dispositivi di I/O, è stata considerata di fondamentale importanza e ha costituito il principale criterio di selezione del campione. A questo fine sono stati seguiti due approcci: uno soggettivo e l'altro oggettivo. A livello soggettivo l'esperienza veniva misurata direttamente tramite una versione abbreviata del questionario EC e indirettamente tramite la versione a 15 item del CATT (§ 5.1.1.2). Il primo questionario è costituito da cinque domande, di cui la prima è l'item di esperienza generale diretta GDE, la seconda e la quarta mirano a elicitare esperienza e conoscenza di linguaggi di programmazione, la terza valuta la frequenza d'uso di applicazioni di videoscrittura³⁹ e la quinta l'abilità d'uso di una tastiera. Il criterio di selezione prescrive un campione caratterizzato da esperienza almeno discreta e per quanto possibile omogenea ($GDE > 2$, media domande 3 e 5 > 2). Inoltre il partecipante doveva dimostrare un atteggiamento verso il computer congruente con l'esperienza dichiarata (valore medio > 3.5 ⁴⁰). La conoscenza di programmazione veniva solo misurata ma non è stata considerata un fattore di selezione nè preventivamente bilanciata

A livello oggettivo la selezione si basava su un veloce test di abilità. Il compito, eseguito su un Macintosh LC, richiedeva di ricopiare nel minor tempo possibile un set di parole differenziate su tre livelli di complessità. Le parole originali erano di volta in volta visualizzate tramite un click del mouse su appositi rettangoli distribuiti sullo schermo. I partecipanti dovevano quindi posizionare il cursore mediante mouse nella porzione inferiore del rettangolo e copiare la parola. Il criterio di selezione si basa su un indice di *efficienza* della prestazione, ottenuto pesando il numero di errori di battitura sul tempo impiegato a eseguire il compito. I risultati del test di abilità confermano la validità del metodo di selezione soggettiva: solo 2 soggetti su 32 (6%) sono stati scartati.

³⁹ Nell'ordine le domande corrispondono agli item 3, 9 e 5 della versione completa del questionario EC.

⁴⁰ Questo valore è stato inferito sulla base dei risultati dello studio 1 e rappresenta il valore medio di atteggiamento dei partecipanti che dichiaravano un livello di esperienza medio.

6.3. Selezione degli attributi di usabilità

L'usabilità è stata misurata sulla base della proposta di Nielsen (§ 3.1) concentrando l'attenzione su tre attributi: *facilità di apprendimento*, *efficienza nell'utilizzo*, *soddisfazione dell'utente*.

Almeno tre caratteristiche del compito e della procedura rendono il concetto di efficacia degli standard ISO poco informativo: (a) il compito è piuttosto facile, di conseguenza non lascia prevedere insuccessi; (b) i partecipanti venivano sollecitati a completare tutti i compiti; (c) le informazioni richieste riguardano reali dati anagrafici, di conseguenza non era possibile verificare la correttezza delle informazioni fornite.

L'attenzione si è dunque concentrata sull'efficienza del sistema. Il costo viene inferito sulla base del numero di scambi richiesti per raggiungere lo scopo. Questo indice sembra infatti più affidabile del *tempo* impiegato a eseguire il compito, che può essere influenzato dall'intervallo necessario alla trasmissione dell'informazione fra i due computer (variabile a seconda del carico di lavoro della rete) e dall'intervallo di reazione del mago (variabile a seconda dell'ambiguità dell'input dell'utente nonché della stanchezza dell'osservatore). Inoltre il criterio di selezione del campione permette di prevedere un tempo di produzione relativamente costante fra i partecipanti. Come suggerito da Nielsen la valutazione dell'efficienza è stata effettuata solo sulla prestazione a regime, una volta terminata la fase di apprendimento. La specifica operazionalizzazione dell'efficienza permette di misurare indirettamente anche la *sicurezza e robustezza all'errore* del sistema: i partecipanti ripetevano un'espressione in seguito a rifiuto del sistema.

A livello dell'efficienza si può prevedere un'interazione IC * Feedback (*ipotesi della flessibilità*). La condizione migliore dovrebbe essere IC Completa e Feedback Presente, la peggiore IC Parziale e Feedback Assente. Nel primo caso infatti flessibilità e controllo sull'andamento dell'interazione sono massimi: la scelta della strategia interattiva è completamente delegata all'utente che ha anche un'immediata risposta dello stato del sistema. A livello di effetti principali invece ci si aspetta solo una generalizzata superiorità della condizione Feedback Presente su Feedback Assente. Infatti, se è vera l'ipotesi del vantaggio il fattore IC non dovrebbe influenzare direttamente la prestazione poiché l'utente dovrebbe essere capace di utilizzare indifferentemente deissi e nominazioni dirette.

La facilità di apprendimento è analizzata invece considerando i dati di tutti i compiti. Una prima verifica riguarda *l'effetto pratica*, operazionalizzato come guadagno in termini di efficienza fra compiti eseguiti in tempi diversi. Ulteriori verifiche riguardano il numero totale di scambi prodotti nei quattro compiti e il comportamento nel primo training. Le ipotesi di lavoro sanciscono innanzitutto la significatività dell'effetto pratica per cui la prestazione dell'utente tende a migliorare con l'esperienza (*ipotesi di adattamento*). Per

quanto riguarda gli specifici effetti dei fattori sperimentali si assume nuovamente l'ipotesi della flessibilità.

La Soddisfazione dell'Utente è stata misurata tramite il questionario SU, uno strumento appositamente realizzato per lo studio e strutturato su due parti. La prima, finalizzata al confronto diretto fra le due condizioni di Feedback, si articola su sei domande. Ogni item richiede di esprimere il giudizio relativo a una specifica dimensione di usabilità su una scala a 7 punti ai cui estremi sono posizionati i due sistemi. I numeri bassi (da 1 a 3) indicano modulazioni della superiorità del sistema dove il feedback è assente, il valore neutro (4) indica uguaglianza, i numeri alti (da 5 a 7) modulazioni della superiorità del sistema dove il feedback è presente. Le dimensioni considerate sono: semplicità d'uso, robustezza all'errore, controllabilità dell'interazione, similitudine con la comunicazione umana, preferibilità, induzione di confusione. La seconda parte invece mira a elicitarne un giudizio generale su MIS mediante 14 item di differenziale semantico, modulati su 7 punti. Gli aggettivi sono stati selezionati a partire da un pool di descrittori utilizzati in precedenti ricerche sull'interazione utente-macchina (Varalda et al., 1994; De Angeli, 1994; De Angeli et al., 1996). Il questionario si chiude con due domande che richiedono un giudizio diretto sulla funzionalità del puntamento.

Relativamente alla soddisfazione dell'utente l'ipotesi principale stabilisce che il giudizio è direttamente correlato alla prestazione (*ipotesi della congruenza*). Di conseguenza ci si aspetta che, congruentemente con l'ipotesi del vantaggio, in generale il sistema venga valutato favorevolmente. I singoli partecipanti dovrebbero poi preferire la condizione di Feedback con cui hanno ottenuto la prestazione migliore. A livello di gruppo, invece, ci si aspetta un vantaggio relativo della condizione IC Parziale in cui il sistema multimodale è effettivamente più utile.

6.4. Metodo

6.4.1. Partecipanti e disegno

Il campione è composto da 30 studenti (15 M e 15 F). Al momento della valutazione l'età media era di 25 anni, da un minimo di 20 a un massimo di 31. Tutti i partecipanti sono di madrelingua italiana e possiedono una discreta esperienza nell'uso di tastiera e mouse (§ 6.2.2).

Il disegno sperimentale prevede la manipolazione del fattore IC (Completa vs. Parziale) fra i soggetti e del fattore Feedback (Presente vs. Assente) entro i soggetti. Le

condizioni di Feedback sono bilanciate fra i partecipanti, mentre l'ordine di presentazione dei compiti è costante.

6.4.2. Procedura

La selezione dei partecipanti veniva effettuata un giorno prima della valutazione. All'inizio della sessione sperimentale gli studenti venivano accolti da uno sperimentatore che li ringraziava per la loro disponibilità e li invitava a leggere accuratamente il foglio informativo sulla valutazione nonché una dettagliata descrizione del funzionamento e delle capacità di MIS differenziata a seconda della condizione di Feedback cui erano assegnati. Secondo la metodologia standard seguita nelle ricerche di interazione utente-macchina, in fase di presentazione veniva posta estrema attenzione nel comunicare ai partecipanti che l'obiettivo della valutazione era il sistema e non le loro caratteristiche comportamentali o di personalità. Inoltre, per evitare di influenzare il comportamento di comunicazione, le istruzioni non fornivano alcun esempio diretto di dialogo. Uno sperimentatore era costantemente disponibile per eventuali chiarimenti.

Dopo il *briefing* iniziale, aveva inizio la valutazione. Conversando con MIS gli studenti dovevano ottenere le istruzioni di compilazione per ogni campo del modulo (domande dell'utente) e fornire i dati personali richiesti (risposte dell'utente). Ogni partecipante compilava due moduli (uno di training e uno sperimentale) per ognuna delle due condizioni di Feedback. Fra una condizione e l'altra era invitato a leggere un foglio di istruzione in cui si specificava il cambiamento dell'interfaccia. Il sistema funzionava secondo le caratteristiche precedentemente descritte.

L'intera sessione sperimentale era videoregistrata da una telecamera a circuito chiuso che collegava la stanza dei maghi a quella dell'utente. Inoltre l'interazione era registrata automaticamente in un file di log. Alla fine i partecipanti compilavano il questionario SU ed erano invitati a riportare le loro impressioni allo sperimentatore in un'intervista libera e informale.

Per evitare di compromettere la credibilità della simulazione i partecipanti sono stati avvisati della vera natura dell'esperimento solo alla fine di tutte le sessioni sperimentali mediante una lettera di spiegazione.

6.5. Risultati

Un primo *screening* del campione ha indotto a eliminare un partecipante della condizione IC Parziale il cui comportamento si discostava significativamente dal campione complessivo e da quello del suo gruppo. In totale dunque sono disponibili i risultati di 58 compiti sperimentali e altrettanti training per un totale di 2.350 espressioni prodotte dagli utenti.

Ai fini delle analisi i log sono stati importati su Excel dove ogni turno conversazionale è stato codificato sulla base di alcune variabili che verranno via via specificate nell'esposizione dei risultati. Il 20% del materiale è stato codificato da due giudici indipendenti. Il confronto dei loro giudizi ha permesso di evidenziare un'affidabilità pari o superiore a .85 per tutte le variabili che verranno analizzate.

6.5.1. Costruzione degli indici soggettivi

La coesione interna della scala di paragone risulta piuttosto soddisfacente ($\alpha = .83$). L'analisi dei valori di correlazione item-totale evidenzia, tuttavia, che il paragone diretto con la comunicazione umana (item 4) non è adeguato a misurare l'efficienza percepita dei due sistemi ($r = -.01$). La sua esclusione permette di aumentare notevolmente il valore di attendibilità ($\alpha = .90$). L'indice di *paragone* è stato quindi calcolato mediando i valori delle rimanenti cinque domande.

Anche la coesione interna della scala di valutazione generale risulta elevata $\alpha = .85$ e tutti gli item contribuiscono al risultato con correlazioni minime superiori a .45. L'indice di *soddisfazione* è stato quindi ottenuto mediando i valori di tutte le coppie di aggettivi, previo ribaltamento di quelle orientate secondo l'ordine positivo-negativo.

Infine è stata misurata la correlazione semplice fra le due domande volte a misurare il giudizio sul puntamento. Il risultato significativo, $r = .61$ ($n = 29$), $p < .001$, permette di mediare i due valori nell'indice di *giudizio sul puntamento*.

6.5.2. Omogeneità del campione

L'omogeneità del campione è stata controllata in riferimento al fattore *IC*, verificando l'equivalenza delle variabili ritenute salienti nei due gruppi sperimentali. A questo fine sono stati calcolati tre indici complessivi. Il primo *abilità I/O* corrisponde alla media delle domande 3 e 5 del questionario di selezione EC; il secondo *atteggiamento* corrisponde alla media degli item del CATT; l'ultimo *programmazione* rappresenta invece la media delle domande 2 e 4. La distribuzione dei valori di tali indici è illustrata nella Figura 14.

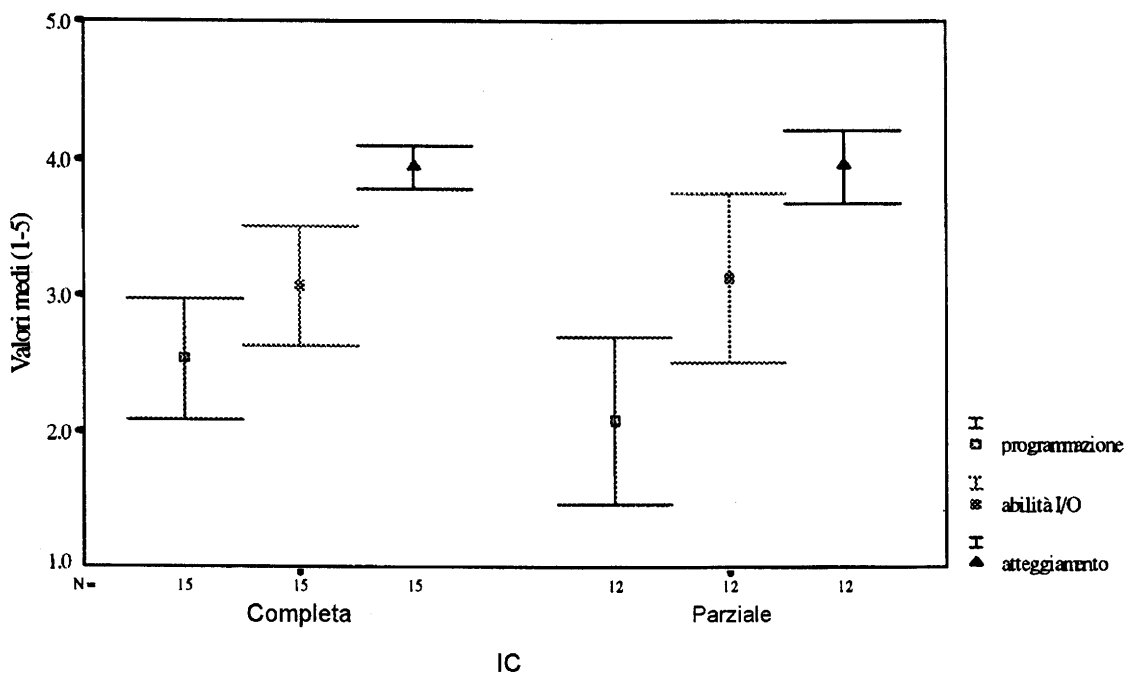


Figura 14. Abilità I/O, atteggiamento e programmazione in funzione di IC.

Le barre verticali corrispondono all'errore standard delle medie N al numero dei casi. Il numero ridotto di casi nella condizione IC Parziale è dovuta a tre missing value nell'indice di atteggiamento.

A causa delle distribuzioni non normali, l'omogeneità è stata valutata mediante tre separati Mann-Whitney U test. I risultati evidenziano effetti non significativi su abilità I/O e atteggiamento [in entrambi i casi $U > .85$, ($n = 29$)], ma tendenzialmente significativi su conoscenza di programmazione [$U(29) = 63.5$, $p = .07$]. Si può quindi concludere che la selezione iniziale e l'assegnamento casuale dei partecipanti nelle due condizioni hanno condotto a un risultato equilibrato relativamente alle variabili di diretto interesse. La lieve disomogeneità sulla conoscenza di programmazione dovrà invece essere tenuta in considerazione nelle prossime analisi.

6.5.3. Ipotesi del vantaggio

Per verificare l'*ipotesi del vantaggio* a livello oggettivo sono stati analizzati i dati dei due compiti sperimentali. Ogni domanda⁴¹ dell'utente è stata codificata a seconda della strategia di identificazione del referente utilizzata. In totale sono state evidenziate quattro diverse strategie: (a) *deittiche*, per esempio "questo -> campo"; (b) *ridondanti*, nominazione più puntamento, per esempio il "campo A ->"; (b) *puramente verbali*, nominazione diretta, indiretta o descrizione linguistica per esempio il "campo A o il campo a destra di A" (c) *puramente gestuali*, puntamenti senza una controparte linguistica. Inoltre ogni domanda è stata codificata come *multimodale* (a+b) o *unimodale*, (c+d). Il grafico in figura Figura 15 riporta le percentuali delle quattro strategie nelle due condizioni di IC.

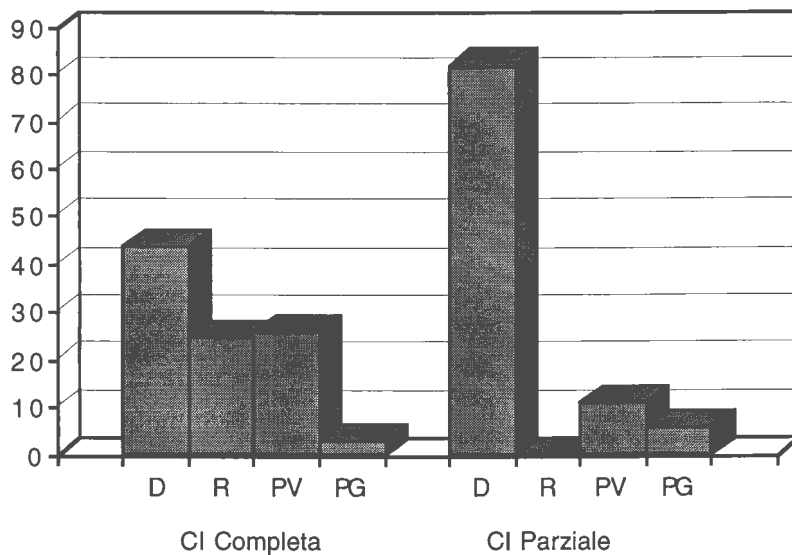


Figura 15. Percentuali delle Strategie di Identificazione del Referente nelle due condizioni di IC.
D=deittiche. R=ridondanti. PV=puramente verbali. PG=puramente gestuale.

Il test di Wilcoxon conferma la maggiore frequenza di strategie deittiche rispetto a quelle puramente verbali sia sul campione completo, $Z = -2.93$ ($n = 29$), $p < .01$, che nella condizione IC Parziale, $Z = -3.01$ ($n = 14$), $p < .01$. Nella condizione IC completa invece le due strategie risultano equiprobabili, $Z = -.94$ ($n = 13$). Comunque il paragone fra frequenza di strategie puramente verbali nelle due condizioni IC non risulta significativo all'analisi del Mann-Whitney U test, $U = 86$ ($n = 29$). La differenza precedente sembra dunque riconducibile all'utilizzo di strategie ridondanti che nella condizione di IC Completa raggiungono il 25% del totale, mentre nella condizione di IC Parziale compaiono una sola volta.

⁴¹ La scelta di effettuare l'analisi solo sulle domande deriva da una caratteristica del modello interattivo del sistema per cui in risposta non era necessario identificare il campo.

Un'ANCOVA sul numero di deissi per il disegno IC (2) * Feedback (2) con numero totale di domande come covariata conferma l'effetto significativo di IC, $F(1, 26) = 6.94, p < .01$, e la non influenza del fattore Feedback né da solo né in interazione, $F(1, 26) < .1$.

In generale dunque gli input multimodali sono utilizzati più frequentemente di quelli unimodali e costituiscono il 76% delle domande, $Z = -3.01 (n = 29), p < .01$.

A livello soggettivo l'ipotesi del *vantaggio* viene confermata mediante più analisi. Innanzitutto si evidenzia che in media la *soddisfazione* si situa nella metà positiva della scala (media = 4.83, SE = .16) ed è significativamente diversa dal valore neutro 4, $t(28) = 5.13 p < .001$. Un'ANOVA su soddisfazione per IC (2) permette di evidenziare che gli utenti delle condizioni Parziale giudicano il sistema un po' più favorevolmente di quelli delle condizioni Completa, $F(1, 23) = 4.57, p < .05^{42}$. Inoltre si evidenzia una correlazione estremamente elevata fra la proporzione deissi/riferimenti totali e l'indice di soddisfazione, $r = .65 (n = 29), p < .001$, nonché proporzione deissi/riferimenti totali e giudizio sul puntamento $r = .632 (n = 29), p < .05 (1 tail)$. Gli utenti che utilizzano più spesso strategie deittiche tendono anche a giudicare più favorevolmente il sistema e a essere più soddisfatti della modalità di puntamento. Queste correlazioni vanno anche a favore dell'*ipotesi della congruenza*.

6.5.4. Ipotesi di produzione

L'ipotesi di produzione è stata verificata dividendo gli input multimodali in due categorie sulla base della posizione del puntamento al loro interno: finale e intermedia. L'analisi del test di Wilcoxon conferma la differenza fra le due posizioni, $Z = -3.03 (n = 24), p < .01$. Come previsto la grande maggioranza (76%) degli input multimodali presenta il puntamento alla fine dello scritto. Il confronto fra posizioni finali sulla base del Feedback conferma un lieve calo nella condizione Presente, $Z = -1.84 (n = 16), p < .05 (1 tail)$. In generale quando il sistema non fornisce alcun feedback lo 81% degli input multimodali presentano il puntamento in posizione finale, mentre quando il campo cambia colore tale percentuale scende al 72%.

⁴² Il numero inferiore di gradi di libertà è dovuto all'esclusione di 3 outlier per rispettare i vincoli imposti dall'analisi della varianza.

6.5.5. Facilità di apprendimento

Consideriamo innanzitutto l'effetto pratica. Poiché i compiti sperimentali richiedevano di compilare un numero maggiore di campi rispetto a quelli di training i valori sono stati uniformati calcolando per ognuna delle quattro prove un *indice di discrepanza* dalla prestazione base. L'indice corrisponde alla differenza fra il numero di scambi prodotti dall'utente e il numero minimo di scambi richiesti per compilare il modulo identificando i campi singolarmente e senza mai commettere errori. Si noti che in questo caso la prestazione base non coincide con la prestazione ideale⁴³, in cui potrebbero venire utilizzate strategie di identificazione multipla.

I quattro indici di discrepanza sono stati quindi analizzati mediante una ANOVA con Posizione Seriale (4) come fattore *within-subject*, IC (2) e Ordine (2) come fattori *between*. Il fattore Ordine corrisponde alla posizione ordinale delle condizioni di Feedback e ha dunque due livelli: Presente-Assente vs. Assente-Presente. Il loro valore permette di riconoscere la condizione di Feedback delle varie Posizioni Seriali.

La Posizione Seriale risulta estremamente significativa, $F(3, 75) = 11.04$, $p < .001$, dimostrando che la prestazione aumenta al crescere dell'esperienza con il sistema. Emerge inoltre la tendenza a un'interessante interazione Posizione Seriale * Ordine, $F(3, 75) = 2.35$, $p = .08$. Da una prima analisi delle medie questo risultato sembra attribuibile alla migliore prestazione della condizione Feedback Equivalente in prima e in ultima posizione. Gli effetti semplici dimostrano comunque che tale superiorità è statisticamente significativa solo in ultima posizione $F(1, 28) = 4.94$ $p < .05$. Come si può osservare nei grafici in Figura 16, quando i soggetti eseguono i compiti nell'Ordine Assente-Presente la prestazione migliora progressivamente. Nel caso di Ordine Presente-Assente invece la curva di apprendimento riproduce una U rovesciata. La prestazione tende a crescere nei primi tre compiti, ma poi ha un'improvvisa caduta sull'ultimo. Il fattore IC non sembra avere nessuna influenza sull'effetto di pratica nè da solo, nè in interazione.

⁴³ Questo è un classico caso di difficoltà di formalizzazione del compito dell'interazione flessibile.

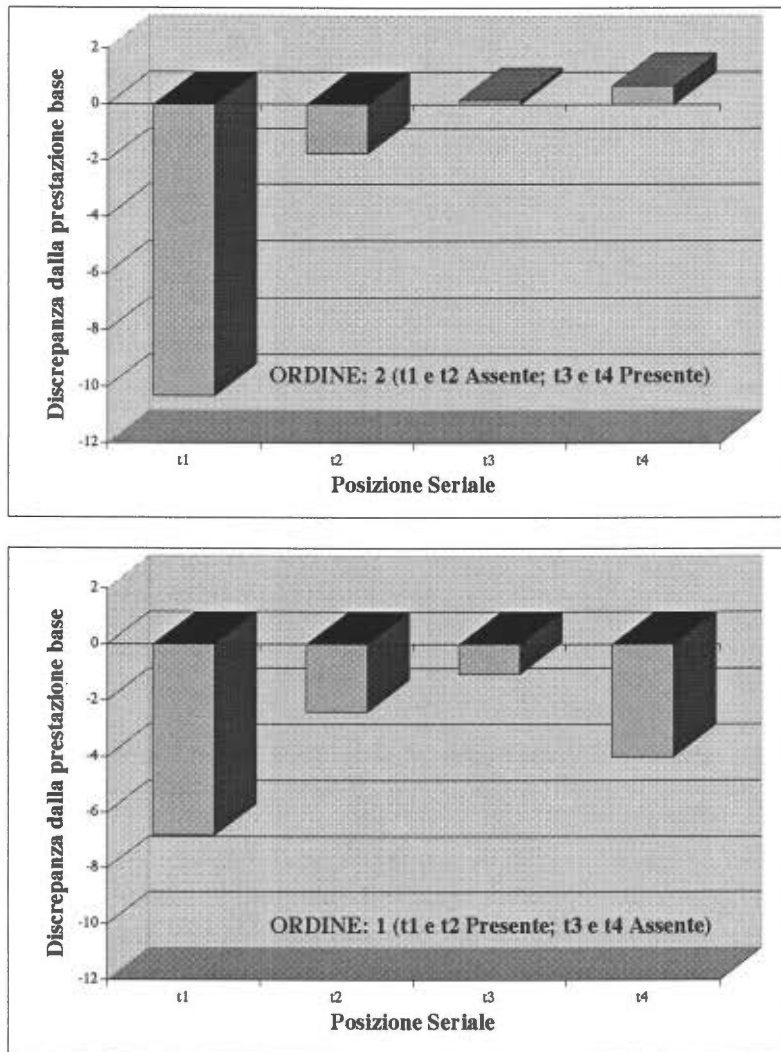


Figura 16. Valori medi degli indici di discrepanza in funzione di Posizione Seriale e Ordine. Lo 0 rappresenta il valore della prestazione base. I valori negativi indicano peggioramento quelli positivi miglioramento.

In sintesi, l'analisi dell'effetto di pratica testimonia che i partecipanti hanno bisogno di tempo per entrare in sintonia con il sistema e raggiungere buone prestazioni, confermando così l'ipotesi di adattamento. La caduta di prestazione nella condizione Ordine Presente-Assente non è semplice da spiegare. Infatti, può essere attribuita a più fattori fra i quali stanchezza degli utenti o un calo di motivazione, che comunque non si verificano nella condizione Ordine Assente-Presente. Inoltre, analizzando mediante una MANOVA i valori di discrepanza della seconda metà della prova (t3 e t4) per il fattore Ordine, si evidenzia un tendenziale ma generalizzato vantaggio della condizione Feedback Presente $F(2, 26) = 2.38, 14.86, p = .11$.

Per valutare fino a che punto si è avuto apprendimento e quando invece la prestazione si è stabilizzata sono state effettuate tre separate analisi della varianza a misure ripetute paragonando i valori di discrepanza di coppie adiacenti di compiti, discrepanza_{t_x} e

discrepanza_{ix+1} con Ordine (2) come fattore *between*. Il passaggio da t1 a t2 indipendentemente dall'Ordine risulta estremamente significativo, $F(1, 27) = 14.86$, $p < .001$. Minore, ma sempre significativa, è anche la differenza fra t2 e t3, $F(1, 27) = 4.63$, $p < .05$. Nuovamente l'ordine non ha influenza, nè come effetto principale, nè nell'interazione. Nel passaggio da t3 a t4 si evidenzia invece una tendenziale interazione Ordine * Posizione Seriale $F(1, 27) = 3.85$, $p = .06$, e l'effetto principale dell'Ordine $F(1, 27) = 4.51$, $p < .05$, ma non della Posizione Seriale in sè, $F(1,27) = 1.94$.

La verifica di differenze così grandi nella prestazione dei partecipanti a seconda della fase della prova induce a ritenere, secondo l'ipotesi di congruenza, che il giudizio soggettivo sul paragone possa a sua volta essere influenzato dall'ordine delle condizioni di Feedback. Un'Anova su paragone per Ordine (2) conferma l'intuizione $F(1, 28) = 5.02$, $p < .05$ e permette inoltre di evidenziare indirettamente una preferenza soggettiva per la condizione Feedback Presente. I partecipanti che hanno usato in fase iniziale il sistema senza Feedback dichiarano poi di preferire il sistema con feedback, mentre quelli che hanno seguito l'ordine opposto li considerano equivalenti. I valori medi e i risultati di un t-test contro il valore neutro quattro sono riportati nella Tabella 8.

	Media	SD	t	gl	2-tail p
Paragone Ordine 1	4.85	.95	4.85	14	.007
Paragone Ordine 2	3.86	1.33	-.41	15	n.s.

Tabella 8. Valori medi di Paragone in funzione di Ordine.

Un'ulteriore analisi si è concentrata sulla verifica delle correlazioni fra l'indice di conoscenza di programmazione e i valori dei quattro indici di discrepanza. L'unico risultato significativo riguarda la Posizione Seriale 1, $r = .41$, ($n = 29$) $p < .05$. Al crescere della conoscenza di un linguaggio di programmazione diminuisce la discrepanza dalla prestazione base, ma solo all'inizio dell'interazione. Anche, l'effetto pratica (discrepanza_{ix}-discrepanza_{ix+1}) risulta correlato alla conoscenza di programmazione, ma nuovamente solo all'inizio del compito, $r = .37$, ($n = 29$), $p < .05$. Questi risultati sono estremamente interessanti in quanto evidenziano l'influenza di conoscenze precedenti sulla prestazione, confermando così l'ipotesi delle differenze individuali che verrà sviluppata e discussa nella prossima simulazione. Per adesso basti considerare che entrambi i compiti sperimentali sembrano scevri dall'influenza di tale fattore.

Dati i risultati ottenuti dall'analisi dell'effetto pratica è necessario domandarsi se sia lecito paragonare i due compiti sperimentali per analizzare l'efficienza. Dato il preventivo bilanciamento delle condizioni di Feedback, l'effetto sequenza non sembra particolarmente preoccupante. Inoltre il paragone diretto fra i due compiti non evidenzia differenze

significative della posizione seriale sulla discrepanza dalla prestazione base, $F(1, 27) < 1$, nè dell'Ordine $F(1, 27) = 2.73, p = .11$ o dell'interazione $F(1, 27) = 2.64, p = .12$.

L'ultima analisi sulla facilità di apprendimento considera un indice generale di prestazione che corrisponde alla somma del numero di scambi prodotti dagli utenti nei quattro compiti. Innanzitutto, si evidenzia un'elevata correlazione fra tale indice e quello di *soddisfazione*, $r = -.53, (n=29), p < .01$. I partecipanti che hanno terminato il compito in un numero minore di scambi sono più propensi a dichiararsi soddisfatti del sistema. Questo risultato conferma l'*ipotesi della congruenza* e sottolinea che i partecipanti valutano la prestazione nel suo insieme, con particolare riferimento alla fase più problematica. Si noti infatti che la correlazione fra soddisfazione e numero di frasi prodotte in t1 è estremamente significativa, $r = -.47, (n=29), p < .01$, mentre la correlazione fra soddisfazione e numero di frasi prodotte in t4 non lo è affatto $r = -.17$.

L'indice generale di prestazione è stato poi inserito come variabile dipendente in un'ANOVA per il disegno IC (2) * Ordine (2). Entrambi i fattori risultano non significativi, $F(1,28) < 1$ sia per gli effetti principali che per l'interazione. La stessa analisi eseguita sul disegno IC (2) * Sesso (2) produce invece un'interessante interazione, $F(1,28) = 4.92 p < .05$. Come si può osservare nel grafico in Figura 17 l'effetto è prevalentemente dovuto alla prestazione delle femmine, che è buona nel caso di IC Completa, mentre peggiora nel caso di IC Parziale. Al contrario, i maschi vanno relativamente meglio nella condizione con meno etichette linguistiche.

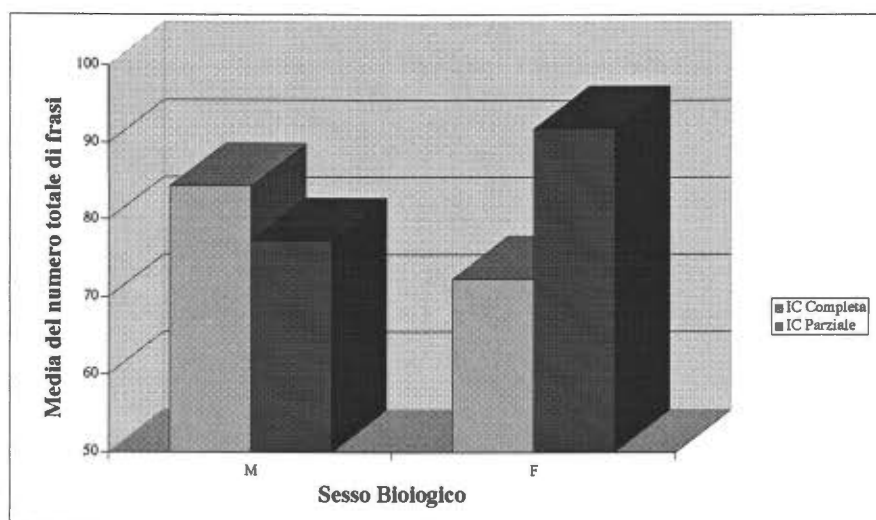


Figura 17. Numero totale di frasi in funzione di IC e Sesso.

Spiegare questa differenza unicamente sulla base delle caratteristiche sessuali può essere difficile e comunque contrario alla prevalente filosofia situazionista sottostante l'approccio globale. Si è così deciso di verificare l'omogeneità delle variabili di selezione sui due sessi. L'analisi, eseguita tramite Mann-Whitney U test, evidenzia una differenza

significativa del Sesso su abilità I/O, $U = 60$, ($n = 29$), $p = .05$. Un'ANOVA sul numero totale di errori ortografici per il disegno Sesso (2) * IC (2) mostra una tendenziale interazione tra i due fattori nello stesso senso di quella evidenziata precedentemente, $F(1,27) = 3.05$, $p = .09$. Nella condizione di IC Parziale le femmine tendono a commettere più errori di ortografia, mentre i maschi ne commettono un po' di più nella condizione IC Completa. Poiché non sono state evidenziate differenze su altre dimensioni di errore, nè sulla strategia di interazione, il gap nella prestazione può essere attribuito principalmente alla differenza nelle abilità I/O.

6.5.6. Efficienza

Una prima verifica dell'efficienza del sistema è stata eseguita tramite una ANOVA a misure ripetute sul numero totale di scambi con Feedback come fattore *within-subject* e IC come fattore *between*. A livello di tendenza emerge sia un'interazione IC * Feedback, $F(1,27) = 3.02$, $p = .09$, sia l'effetto principale del Feedback, $F(1,27) = 2.60$, $p = .12$. Il fattore IC invece risulta del tutto estraneo $F(1,27) < 1$. Come previsto dall'*ipotesi della flessibilità* i partecipanti assegnati alla condizione IC Completa e Feedback Presente completano il modulo con un numero medio di frasi inferiore a quello richiesto dalle altre condizioni. Come si può vedere nel grafico in Figura 18 l'interazione è completamente spiegata da questo effetto, le altre condizioni hanno un comportamento omogeneo.

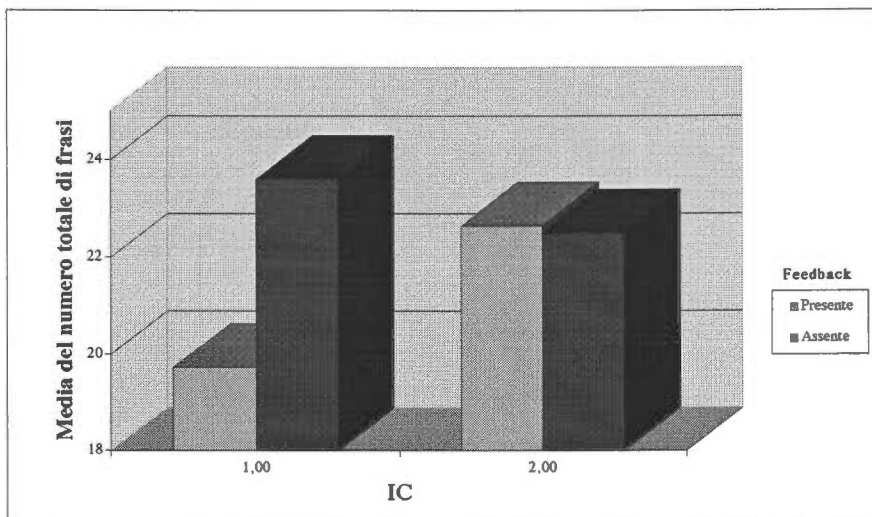


Figura 18. Numero medio di frasi nei due compiti in funzione di IC e Feedback

Per diminuire il possibile effetto dovuto a stili interattivi diversi, più o meno propensi a commenti che esulano dal compito, è stata effettuata una seconda analisi sullo stesso disegno della precedente considerando solo il numero di domande, previa esclusione di due casi che si allontanavano in modo significativo dalla media della distribuzione. L'interazione raggiunge la piena significatività statistica $F(1,25) = 6.70$, $p < .05$ e cresce anche l'effetto relativo del Feedback $F(1,25) = 3.24$, $p = .08$. Nuovamente le prestazioni migliori sono ottenute dal gruppo IC Completa e Feedback Presente. In media gli utenti completano il modulo con circa 9 domande. La condizione IC Parziale e Feedback Assente è invece molto simile alle altre. Possiamo dunque considerare parzialmente confermata l'ipotesi della flessibilità.

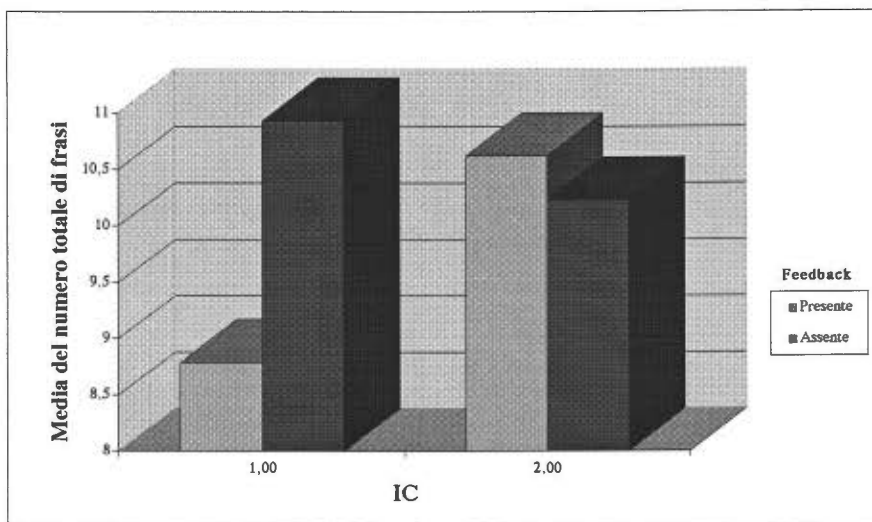


Figura 19. Numero medio di domande in funzione di IC e Feedback

6.6. Conclusioni

A conclusione di questo primo studio possiamo fare le seguenti osservazioni. A livello metodologico si conferma l'utilità di selezionare i partecipanti prima del loro assegnamento casuale alle condizioni sperimentali. Il criterio di selezione soggettivo mediante misurazione incrociata di esperienza e atteggiamento risulta estremamente efficiente. Il controllo oggettivo ha respinto solo il 6% dei partecipanti ritenuti idonei sulla base del criterio soggettivo.

In generale l'approccio seguito nella valutazione è risultato soddisfacente. L'analisi di varie misure eterogenee ha permesso di evidenziare un *trend* di risultati senza incongruenze. L'ipotesi della *congruenza* è dunque ampiamente supportata dai risultati. Si noti però che la valutazione soggettiva si basa sul giudizio complessivo della prestazione. In altre parole i partecipanti tendono a giudicare la prestazione generale piuttosto che l'interazione a regime. Di conseguenza la fase di training può essere utile a uniformare la prestazione del sistema interattivo su livelli di efficienza più o meno stabili, ma non serve a cancellare dalla memoria dell'utente le difficoltà iniziali, su cui egli basa prevalentemente il giudizio. Queste considerazioni possono aiutare a spiegare la mancanza di correlazione spesso evidenziata in precedenti studi (§).

L'*ipotesi del vantaggio* risulta confermata sia a livello soggettivo che oggettivo. È stato infatti provato che i sistemi multimodali sono utili e piacciono anche quando impongono procedure non naturali. Ciò è dovuto, secondo me, al grande valore e alla potenza della comunicazione non verbale (McNeill, 1992). Comunque è stato dimostrato che anche i sistemi flessibili necessitano di un periodo di '*rodaggio*' per riuscire a raggiungere prestazioni soddisfacenti, confermando così l'*ipotesi dell'adattamento*. Ciò testimonia che l'interazione rimane vincolata a criteri diversi da quelli che guidano la comunicazione fra umani. La differenza fra *conversazione* e *interazione* è inoltre testimoniata dal frequente uso di strategie di identificazione ridondanti, evidenziato nella situazione di Identificabilità dei Campi Completa. Tale soluzione è sintomo di una tendenza a un'eccessiva precisione nell'interazione, mentre la comunicazione fra umani è spesso caratterizzata dalla tendenza opposta.

L'ipotesi della *flessibilità* è parzialmente validata. La situazione più ergonomica, in cui l'utente è realmente libero di scegliere la strategia preferita e ha un elevato controllo sull'interazione, risulta anche la situazione più efficiente. Comunque non si manifestano differenze fra le altre condizioni. Possiamo dunque concludere, congruentemente con l'approccio globale, che l'usabilità di un sistema dipende dal suo insieme e non da singole parti separate.

In generale, la presenza del Feedback tende a migliorare la prestazione e rende la formulazione della deissi meno vincolata alla sequenza scrittura - puntamento. Inoltre è la soluzione preferita dagli utenti. A questo proposito è utile aprire una breve parentesi per alcune considerazioni sulla posizione sul continuum naturale-artificiale delle due condizioni di Feedback. I fautori dell'assenza del Feedback motivano la loro scelta affermando una maggiore similitudine con la comunicazione umana (Schmauks, 1987). Eppure, se è vero che, quando si indica un oggetto il target in sé non subisce alcuna modifica fisica, è anche vero che il parlante ha a disposizione immediato feedback da parte dell'ascoltatore. Quest'ultimo infatti sposta lo sguardo nella direzione dell'oggetto indicato fornendo così un veloce e generale segnale relativo alla sua comprensione. In questo senso, dunque, ogni azione di puntamento naturale è correlata a un feedback da parte dell'ascoltatore. Dunque entrambe le soluzioni mimano la comunicazione umana. La presenza del feedback in qualche modo la rende più potente, aumentando la forza di un comportamento naturale. Comunque, secondo me, l'interazione utente-computer deve svilupparsi sulla base delle sue peculiari caratteristiche e non cercare di ricopiare passivamente la comunicazione faccia a faccia.

7. Simulazione 2

L'esperimento di simulazione descritto in questo capitolo si propone di indagare nel dettaglio uno dei risultati evidenziati dal precedente lavoro, vale a dire la correlazione fra conoscenze di programmazione e prestazione iniziale durante l'interazione. Dunque, l'interesse si orienta prevalentemente sull'analisi dell'effetto di una specifica determinante relativa all'utente sul comportamento di interazione di sistemi più o meno flessibili.

La metodologia seguita nei due studi è piuttosto simile, di conseguenza verranno esclusivamente specificate eventuali differenze, per il resto si rimanda alla dettagliata esposizione fornita nel capitolo 6.

7.1. Definizione degli obiettivi

Lo studio costituisce la valutazione comparata tramite simulazione semiautomatica di due sistemi, identificati dall'acronimo SIM (*Sistema Intelligente Multimodale*). I sistemi si differenziano per il loro livello di flessibilità. Il primo è unimodale e supporta l'interazione tramite linguaggio scritto, l'altro è multimodale e supporta l'interazione tramite linguaggio scritto e puntamento. I sistemi sono stati valutati da due gruppi di utenti caratterizzati da un diverso loro livello di esperienza informatica, operazionalizzata come familiarità con il computer e conoscenza di programmazione. Una parte del campione è costituita da utenti esperti nell'uso del computer e con discrete conoscenze di programmazione, l'altra da persone che non avevano mai o quasi mai usato il computer.

Il compito era abbastanza simile al precedente con la differenza però che agli utenti veniva dato un limite temporale. Dopo cinquanta minuti SIM si spegneva automaticamente. Nel tempo a loro disposizione i partecipanti dovevano compilare due moduli. Il primo (Modulo A) riproduce quelli usati nella simulazione 1, ma tutti i campi sono chiaramente identificati da scritte, per esempio *dati anagrafici*. Il secondo (Modulo B) invece è più complesso e richiede di eseguire sia azioni di puntamento che di spostamento, tipo "*metti (target A) in (target B)*".

Le ipotesi di lavoro derivano dalla combinazione di quelle descritte precedentemente (§ 6).

Innanzitutto ci si aspetta che l'esperienza dell'utente tenda a influenzare la prestazione, ma solo nel primo periodo di adattamento (ipotesi della *modulazione delle differenze individuali*). Poi si ipotizza un'interazione Sistema * Modulo per cui il vantaggio dell'interazione multimodale dovrebbe emergere solo nel modulo B (ipotesi della *modulazione del vantaggio*). Il modulo A infatti permette di identificare i campi

molto semplicemente tramite nominazione diretta e, come dimostrato dalla simulazione 1, le due strategie conducono a risultati molto simili.

Riassumendo, le aspettative sono congruenti con il seguente pattern: durante la fase di apprendimento(modulo A) gli utenti esperti dovrebbero andare meglio degli inesperti, indipendentemente dal sistema utilizzato che non dovrebbe influenzare la prestazione. Col tempo (modulo B) il comportamento di esperti e inesperti dovrebbe coincidere e la peculiarità del compito dovrebbe fare emergere chiaramente il vantaggio del sistema multimodale.

7.2. *Caratterizzazione delle entità*

La caratterizzazione dell'ambiente è identica a quella dello studio precedente, mentre, come già accennato, il compito è un po' diverso. Innanzitutto ai partecipanti viene consegnato uno scenario dove sono descritte tutte le informazioni da inserire nei moduli. La compilazione prevede dunque soluzioni esatte. L'obiettivo dell'utente è compilare due moduli, impersonificando uno studente che vuole recarsi in Olanda con una borsa di studio di un anno.

Il modulo A richiede di inserire i dati socio-anagrafici elencati nello scenario secondo i criteri di compilazione richiesti da SIM. La compilazione del modulo B è invece più impegnativa. La procedura di compilazione è parzialmente diretta dal compito: alcune sezioni devono essere riempite prima di altre. Scopo del compito è preparare un piano di studio scegliendo fra una lista di esami prestampata sul modulo. I vincoli di selezione sono esplicitati nello scenario e gli esami sono scritti in olandese. Durante l'interazione il partecipante deve dunque chiedere a SIM la traduzione degli esami e alcune informazioni generali. Sulla base di tali conoscenze deve poi decidere se l'esame è o meno adeguato ai vincoli richiesti.

Il modello interattivo del sistema è in generale un po' più flessibile del precedente. SIM infatti accetta le ellissi e in caso di incomprensione fornisce una specificazione del problema comunicativo con frasi tipo "*Mi dispiace, la sua frase è troppo complessa per me. Provi a semplificarla*" oppure "*Mi dispiace, non conosco la parola [parola target]. Provi a usare un sinonimo*". Inoltre è più cooperativo e direttivo.

Il campione è stato selezionato esclusivamente sulla base di auto-valutazioni soggettive (questionario CATT e una forma adattata del questionario EC).

L'analisi dei dati si concentra prevalentemente sulla facilità di apprendimento. Di conseguenza vengono riportati solamente i giudizi soggettivi relativi a tale dimensione. Le opinioni dei soggetti sono state misurate prima e dopo l'utilizzo del sistema mediante due questionari che riproponevano con un ordine diverso le stesse scale di valutazione. Di conseguenza sono disponibili un indice di *semplicità attesa* e uno di *semplicità effettiva*.

7.3. Metodo

7.3.1. Partecipanti e disegno

Il campione è composto da 50 studenti universitari di madre lingua italiana (28 M e 20 F) selezionati secondo il criterio descritto precedentemente. Al momento della valutazione, avvenuta nell'ottobre del 1996, l'età media era di 23 anni, da un minimo di 19 a un massimo di 34. Tutti i partecipanti sono stati pagati per la loro collaborazione.

Il disegno sperimentale prevede un disegno misto 2*2*2 con Sistema (unimodale vs. multimodale) ed Esperienza (esperti vs. inesperti) manipolati fra i soggetti, Compito (modulo A vs. modulo B) manipolato entro i soggetti. La sequenza di compilazione dei moduli era costante.

7.3.2. Procedura

La selezione dei partecipanti veniva effettuata un giorno prima della valutazione. All'inizio della sessione sperimentale lo sperimentatore consegnava agli studenti un foglio informativo sulla valutazione e una dettagliata descrizione del funzionamento e delle capacità di SIM differenziata a seconda del sistema con cui avrebbero dovuto interagire. In fase di presentazione veniva posta estrema attenzione nel chiarire che l'obiettivo della valutazione era la capacità interattiva del sistema non la prestazione o le caratteristiche di personalità degli utenti. Le istruzioni erano elaborate in modo da non includere nessun esempio diretto di dialogo.

Dopo il *briefing* iniziale ai partecipanti veniva chiesto di compilare il questionario SU1, poi aveva inizio la valutazione. Conversando con SIM gli studenti dovevano ottenere le istruzioni di compilazione dei moduli (domande dell'utente) e fornire i dati richiesti (risposte dell'utente). Ogni partecipante aveva a disposizione 50 minuti e due moduli da compilare (modulo A e modulo B). Il sistema funzionava secondo le caratteristiche precedentemente descritte.

L'intera sessione sperimentale era registrata da telecamera e da computer. Dalla stanza di controllo due maghi monitoravano il comportamento dell'utente. Alla fine i partecipanti compilavano il questionario SU2 ed erano sollecitati a riportare le loro impressioni in un'intervista libera e informale.

Come nella simulazione precedente la vera natura dell'esperimento è stata comunicata a tutti i partecipanti contemporaneamente tramite una lettera inviata alla fine di tutte le sessioni.

7.4. Risultati

Un primo *screening* del campione ha indotto a eliminare due partecipanti, uno per problemi informatici relativi al file di log, l'altro perché la sua prestazione si differenziava significativamente da quella di tutti gli altri utenti. Dopo tale selezione i gruppi sperimentali risultano composti da 12 persone ciascuno.

I log delle interazioni sono stati importati in Excel per un'analisi dettagliata del comportamento. L'indice di attendibilità ottenuto dalle valutazioni di due giudici indipendenti sul 20% del materiale ha fornito risultati soddisfacenti per tutte le variabili dipendenti analizzate.

7.4.1. Verifica delle ipotesi sperimentali

La prima verifica ha riguardato il paragone del tempo impiegato per compilare il modulo A nelle diverse condizioni sperimentali. La variabile, normalizzata tramite calcolo della radice quadrata, è stata quindi sottoposta a una ANOVA per il disegno Sistema (2) * Esperienza (2). I risultati evidenziano unicamente un effetto significativo dell'Esperienza $F(1, 47) = 23.60, p < .001$. Come previsto dall'ipotesi delle *differenze individuali*, i partecipanti più esperti, hanno completato la prova in un tempo significativamente inferiore dei meno esperti. La mancanza di significatività del fattore Sistema è congruente con l'ipotesi della *modulazione del vantaggio*.

Essenzialmente questo risultato può essere attribuito a varie cause: migliore comprensione del compito, maggiore velocità di adattamento linguistico e maggiore velocità di produzione degli input. Quest'ultima può essere ragionevolmente ritenuta valida, ma non è stato possibile sottoporla a verifica diretta. Dal log del sistema non si riesce infatti a distinguere l'effettivo tempo di delle parole. Le prime due ipotesi invece sono state sottoposte a verifica.

L'ipotesi della comprensione del compito è stata testata confrontando il numero di frasi non direttamente correlate al compito di compilazione prodotte dai due gruppi di diversa esperienza. La differenza risulta significativa all'analisi del Mann-Whitney U test, $U(48) = 193.0, p < .5$. Nella prima fase dell'interazione il comportamento dei meno esperti è maggiormente caratterizzato da domande di chiarimento sulle capacità del sistema e del compito da eseguire, da generiche richieste di aiuto nonché da frasi di mantenimento del dialogo. Tali espressioni costituiscono il 13% del numero totale di scambi nel campione dei meno esperti e il 8% in quello degli esperti.

La verifica dell'ipotesi dell'adattamento linguistico si è concentrata invece sul confronto del numero di errori comunicativi. La variabile è stata analizzata tramite ANOVA sul disegno Sistema (2) * Modulo (2). Come evidenziato dal grafico in Figura

20 i risultati mostrano che le persone meno esperte fanno più errori di quelle esperte $F(1, 47) = 4.93, p < .05$. Non si evidenziano invece effetti del Sistema.

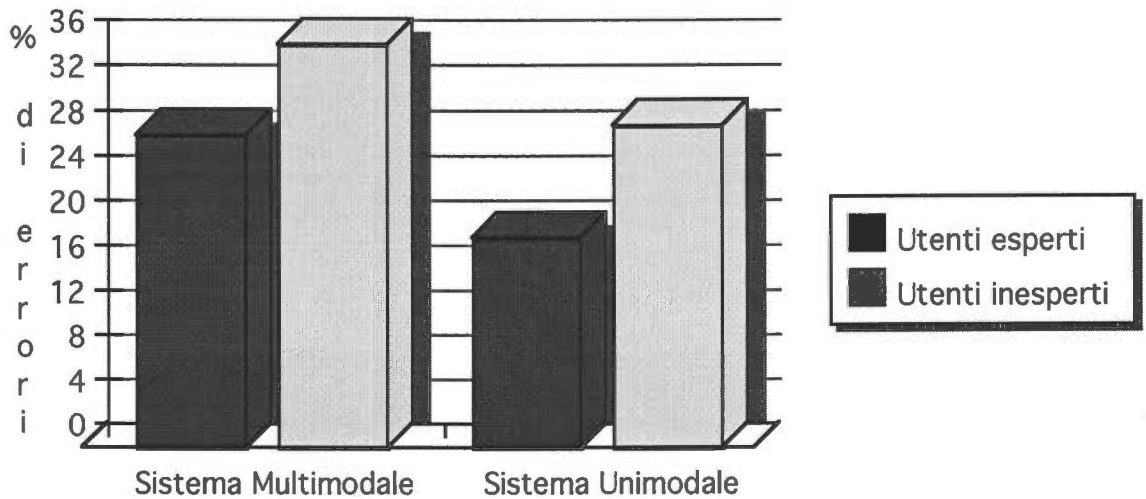


Figura 20. Percentuale di errori comunicativi nelle condizioni sperimentali .

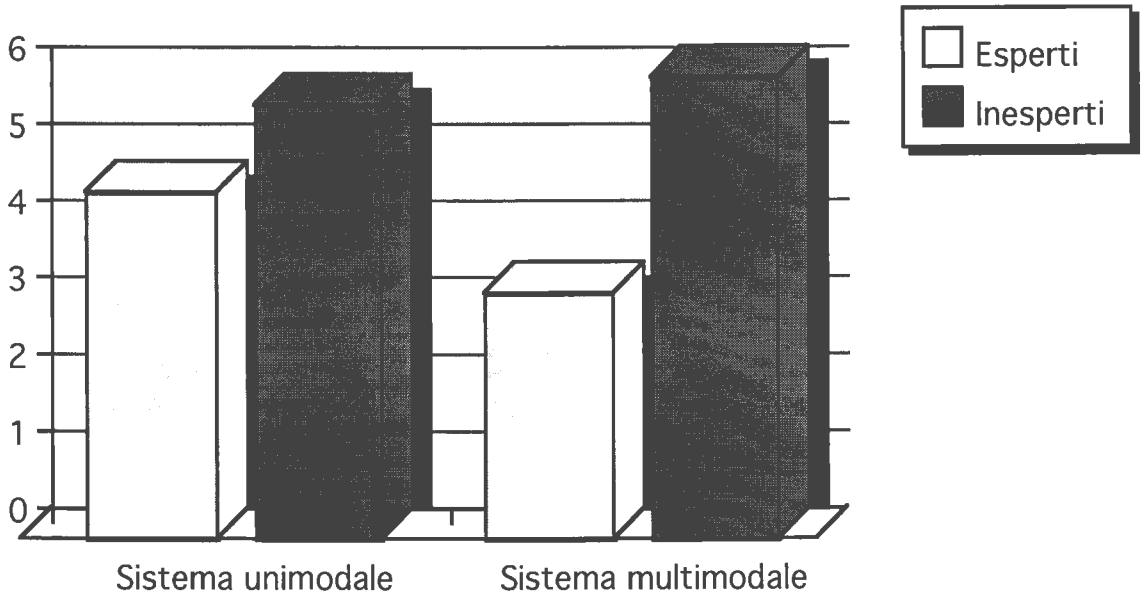
Concludendo la migliore prestazione dei programmatori può essere attribuita sia a una più veloce comprensione del compito che a una migliore prestazione comunicativa.

Il confronto fra il modulo A e il modulo B è stato effettuato sull'indice di lunghezza media della frase. La validità dell'indice, ottenuto dividendo il numero di parole scritte dall'utente per il numero di frasi, è confermata dalla letteratura e da due analisi di correlazione effettuate sui dati in esame. Tali verifiche hanno premesso di evidenziare che i partecipanti che usano frasi meno lunghe, tendono anche a totalizzare meno fallimenti comunicativi, $r = .48, (n = 48) p < .01$ e a produrre un maggior numero di input multimodali, $r = -.47, (n = 24) p < .01$.

Entrambi gli indici sono stati normalizzati mediante trasformazione logaritmica e paragonate tramite una ANOVA a misure ripetute con Sistema e Esperienza come fattori between-subject. L'analisi evidenzia la significatività dell'interazione Sistema * Compito $F(1, 40) = 8.71, p < .001$, nonché gli effetti principali dell'Esperienza $F(1, 40) = 5.01, p < .05$ e del Modulo $F(1, 40) = 4.93, p < .05$. Come si può osservare nei grafici in Figura 21 l'elevata significatività dell'interazione può essere attribuita al notevole accorciarsi delle frasi nella condizione multimodale e al lieve allungarsi delle stesse nella condizione Unimodale in corrispondenza al Compito B. Quest'ultimo effetto è attribuibile alla maggiore complessità di identificazione dei campi del modulo B. Il primo invece, come evidenziato da un Wilcoxon-test, dipende da un generalizzato aumento del numero di

input multimodali, $Z = 3.74$ ($n = .21$), $p < .001$, che nel primo compito rappresentano il 38% delle strategie di identificazione e nel secondo il 71%.

Modulo A



Modulo B

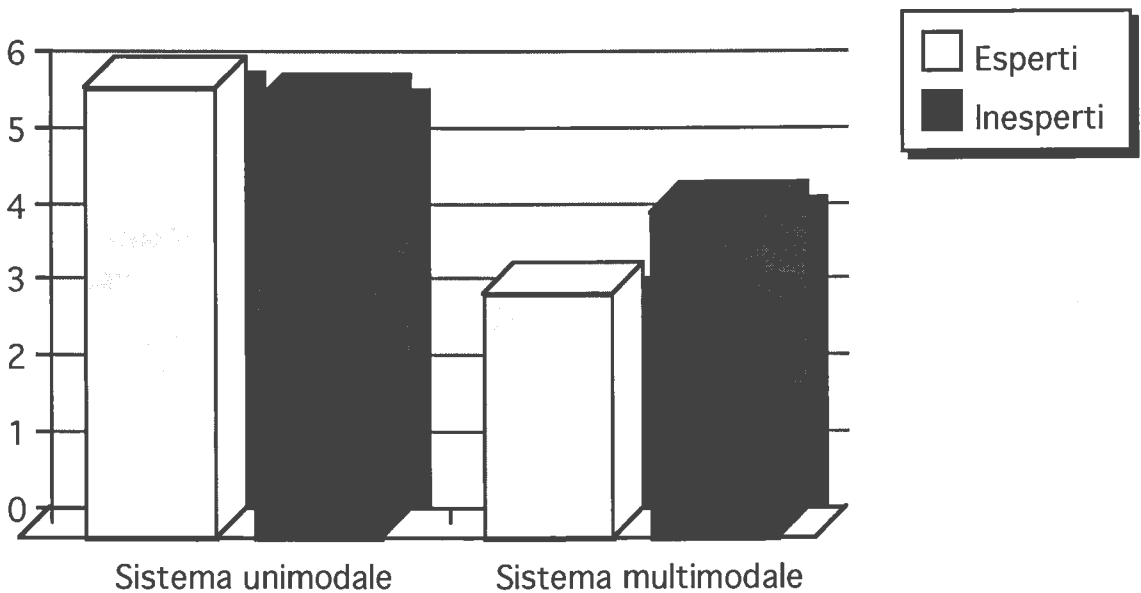


Figura 21. Lunghezza delle frasi in funzione di Modulo e Sistema.

L'analisi degli effetti semplici permette di chiarire che l'effetto dell'Esperienza è prevalentemente dovuto alla prestazione nel Compito A, $F(1, 40) = 7.95$, $p < .01$ e tende a scomparire nel Compito successivo $F(1, 40) = 1.07$.

L'analisi dei giudizi soggettivi sull'indice di semplicità attesa e semplicità effettiva è congruente con il pattern di risultati oggettivi. Da una ANOVA a misure ripetute per il disegno Esperienza * Sistema emerge un chiaro effetto dell'Esperienza sul giudizio, $F(1, 44) = 5.82, p < .05$. Gli utenti meno esperti si aspettano e valutano l'interazione come più complessa.

7.5. Conclusioni

Entrambe le ipotesi sperimentali sono state confermate. L'esperienza discrimina la prestazione ma solo all'inizio, durante il periodo di adattamento. Il vantaggio dei più esperti è dovuto sia a una loro maggior velocità di comprensione del compito e delle strategie di soluzione permesse dal sistema informatico che a un più veloce adattamento linguistico.

Il vantaggio dei sistemi multimodali rispetto agli unimodali emerge esclusivamente in compiti per cui il linguaggio è poco efficiente.

7.6. *Discussione finale*

Le ricerche sperimentali descritte in questa tesi permettono di sostenere che i sistemi multimodali sono potenti strumenti di interazione. Il loro valore deriva dalla flessibilità dialogica, conseguente alla disponibilità di più canali di comunicazione. In primo luogo, dunque, il vantaggio di tali sistemi è funzione del significato trasmesso. Gli input multimodali risultano particolarmente efficienti per identificare oggetti nello spazio, specialmente nei casi in cui la nominazione diretta non è possibile (Bellalem et al., 1996; Oviatt, 1996; Oviatt et al, 1997).

La potenza dei sistemi flessibili è testimoniata dal veloce coincidere della prestazione di utenti esperti e novizi, inimmaginabile nell'interazione con sistemi vincolati. Il vantaggio relativo ai sistemi unimodali di comprensione del linguaggio consiste nell-induzione di un dialogo più appropriato alla macchina. In generale l'utente tende a produrre frasi più brevi e a diminuire gli errori comunicativi. Di conseguenza dunque aumentano le possibilità di riconoscimento.

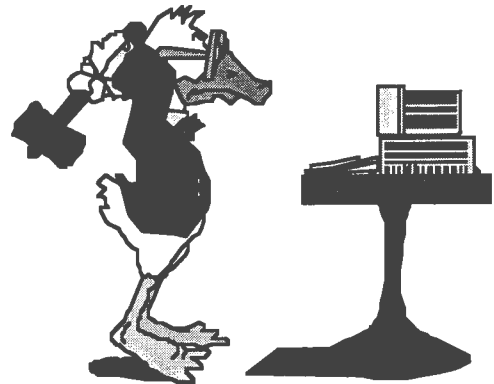
Eppure il divario fra utenti e computer rimane anche quando l'interazione riproduce da vicino la comunicazione fra umani. È stato infatti dimostrato che gli attuali sistemi multimodali non permettono un immediato accesso alla tecnologia. L'efficacia nella prestazione è ancora subordinata a una fase di training in cui l'utente deve comprendere le capacità interattive del computer e adattarvi il suo comportamento. Questa fase è soggetta a notevoli differenze che sembrano dipendere più da specifiche caratteristiche dell'utente che del sistema.

Gli esperimenti discussi dimostrano l'importanza dell'esperienza informatica pregressa sul periodo di adattamento. Gli utenti maggiormente esperti riescono ad adeguare il loro comportamento più velocemente di quelli meno esperti. Probabilmente, gli esperti applicano le loro conoscenze passate, relative a sistemi informatici vincolati, al nuovo tipo di interazione. In qualche modo dunque sottostimano e sottovalutano il sistema, ma sono più efficienti.

E' evidente che l'evoluzione della dimensione cognitiva delle interfacce dovrebbe accelerare il processo di adattamento rendendolo biunivoco. Fino a quel momento però è necessario comunicare all'utente in modo adeguato e non ambiguo le effettive capacità di comprensione del sistema per evitare di generare aspettative che non possono essere poi rispettate. La delusione potrebbe infatti indurre un processo di rifiuto della tecnologia, specialmente da parte degli utenti meno esperti a cui viene richiesto il maggiore sforzo di adattamento. Infatti, contrariamente a quanto dimostrato in altri studi (§ 4.4) gli utenti si dimostrano giudici sensibili dell'interazione. La congruenza la loro valutazione e le misure standard di efficacia è però mediata da vari fattori che devono essere tenuti in

considerazione nell'analisi. Il giudizio soggettivo si basa, infatti, sulla prestazione completa e non su una singola parte come tradizionalmente avviene nella valutazione dell'efficacia (Nielsen, 1993).

A livello generale dunque i lavori sperimentali presentati in questa tesi dimostrano la difficoltà della ricerca empirica nello studio dei sistemi interattivi flessibili. L'adeguatezza della prospettiva teorica e metodologica dell'approccio globale sembra confermata dalla sua capacità di fornire spiegazioni congruenti dell'oggetto di studio e dalla sensibilità discriminativa dimostrata nella prova sul campo. Certo l'approccio deve ancora essere raffinato e migliorato, ma dimostra potenzialità interessanti.



Riferimenti Bibliografici

- Abbagnano, N. (1982). *Storia della Filosofia*. Torino: UTET, vol 1.
- ACM SIGCHI, (1992). *Curriculum for Human-Computer Intercation*. ACM Special Interest Group on Computer-Human Interaction Curriculum Development Group, New York.
- Airenti, G., Bara, B. G. e Colombetti, M. (1987). Modelli della comunicazione uomo-macchina. In O. Stock (a cura di) *L'interazione uomo-computer*, Roma: Il Pensiero Scientifico Editore, pp. 137-156.
- Anolli, L. e Ciceri, R. (1995). *Elementi di psicologia della comunicazione*. Milano: LED edizioni universitarie di Lettere Economia e Diritto.
- Archer, N. P. e Yuan, Y. (1995). Comparing telephone-computer interface designs: Are software simulations as good as hardware prototypes?. *International Journal of Human-Computer Studies*, 42, pp.169-184.
- Arcuri, L. (1985). *Conoscenza sociale e processi psicologici*. Bologna: Il Mulino.
- Arens, Y. e Hovy, E. (1990). How to Describe What? Towards a Theory of Modality Utilization. *Proceedings of the 12th Annual Conference of the Cognitive Science Society*, July.
- Baecker, R. M. e Buxton, W. A. S. (1987). The Context of Human-Computer Interaction. In R. M. Baecker e W. A. S. Buxton (a cura di) *Readings in Human Computer Interaction*. San Mateo CA:Morgan Kaufmann.
- Baecker, R. M., Small, I. e Mander, R. (1991). Bringing icons to live. *Proceedings ACM CHI'91 Conference*, New Orleans, LA, 28 April-2 May, pp. 1-6.
- Bagnara, S. e Visciola, M. (1987). Ergonomia e interazione uomo-macchina a base informatica. In O. Stock (a cura di) *L'interazione uomo-computer*, Roma: Il Pensiero Scientifico Editore, pp. 19-31.
- Bailey, J. E. e Pearson, S. W. (1983). Development of a tool for measuring and analyzing computer user satisfaction. *Management Science*, 29, pp. 530-545.
- Bandalos, D. e Benson, J. (1990). Testing the factor structure invariance of a computer attitude scale over two grouping conditions. *Educational and Psychological Measurement*, 50, pp. 49-60.
- Bannon, L. (1991). From human factors to human actors: the role of psychology and human-computer interaction studies in system design. In J. Greenbaum e M. King (a cura di) *Design at work : cooperative design of computer systems*, Hillsdale, NJ: Lawrence Erlbaum Associates
- Bara, B. G. (1984). Artificiale e No. *Storia e Critica della Psicologia*, 1, pp. 43-57.
- Baroudi, J. J., Olson, M. H. e Ives; B. (1986). An Empirical Study of the Impact of user Involvement on System Usage and Information Satisfaction. *Communication of the ACM*, 29, pp. 231-238.
- Beimel, J., Schindler, R. e wandke, H. (1994). Do human factors expert accept the ISO 9241 Part 10-Dialogue Principle-standard?. *Behaviour & Information technology*, 13, pp. 299-308.
- Bellale, N., Romary, L. e Wolff, F. (1996). Gestural prosody in man-machine task-oriented dialogues: some hypohese. *Proceedings of the Workshop on the Integration of Gesture in Language and Speech*, October 7-8, Newark and Wilmington Delaware, pp. 205-213

- Bem, S. L. (1974). The measurement of psychological androgyny. *Journal of Consulting and Clinical Psychology*, 42, pp. 155-162.
- Bem, S. L. (1981). *Bem Sex-Role Inventory: Professional Manual*. Palo Alto: Consulting Psychologist Press.
- Benysh, D. V., Koubek, R. J. e Calvez, V. (1993). A comparative Review of Knowledge Structure Measurement Techniques for Interface Design. *International Journal of Human-Computer Interaction*, 5, pp. 211-237.
- Blattner, M. M., Sumikawa, D. A. e Greenberg, R. M. (1989). Earcons and Icons: Their Structure and Common Design Principles. *Human-Computer Interaction*, 4, pag 11-44
- Bodker, S. (1989). A Human Activity Approach to User Interface. *Human-Computer Interaction*, 4, pp. 171-195.
- Booth, P. (1989). *An Introduction to Human-Computer Interaction*. Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- Borchers, J., Deussen, O. e Knorz, C. (1995). *SIGCHI Bulletin*, 27, pp. 72-78
- Brighetti, G., Contento, S. e Tabossi, P. (1987). I modelli dell'utente nell'interazione uomo-computer. In O. Stock (a cura di) *L'interazione uomo-computer*, Roma: Il Pensiero Scientifico Editore, pp. 33-50.
- Brouwer-Janse, M. D. (1991). Users are problem solvers! *IPO Annual Progress Report*, 26, pp. 113-121.
- Bubb, P., De Angeli, A., Ferrante, D., Haller, R., Jacomussi, P., Toffetti, A. e Varalda, G. (1994). *Device Standards and Extensions of Regulations for ATT Input and Dialogue Devices*. Deliverable n° 13, Type: Public, Drive II Project V2006 EMMIS, September.
- Campbell, R. L e Bickhard, M. H. (1986). *Knowing levels and developmental stages*. Basel: S. Karger.
- Card, S. K., Moran, T. P. e Newell, A. (1980). The Keystroke-Level Model for User Performance Time with interactive Systems. *Communication of the ACM*, 23(7), pp. 396-410.
- Card, S. K., Moran, T. P. e Newell, A. (1983) *The psychology of human computer interaction*. Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- Carroll, J. M. , Kellog, W. A. e Rosson, M. B. (1991). The Task-Artifact Cycle. In J. M. Carroll (A cura di) *Designing Interaction: Psychology at the Human-Computer Interface*. New York: Cambridge University Press.
- Carroll, J. M. a cura di (1987). *Interfacing Thought: Cognitive Aspects of Human-Computer Interaction*. Cambridge, MA: MIT Press.
- Carroll, J. M. e Campbell, R. L. (1986). Softening Up Hard Science: Reply to Newell and Card. *Human-Computer Interaction*, 2, pp. 227-249.
- Carroll, J. M. e Olson, J. R. (1988). Mental Models in Human-Computer Interaction. In M. Helander (a cura di) *Handbook of Human-Computer Interaction*, Amsterdam: North-Holland, pp. 45-85.
- Carroll, J. M., Kellog, W. A. e Rosson, M. R. (1991). The Task-Artifact Cycle. In J. M. Carroll (a cura di) *Designing Interaction: psychology at the Human-Computer Interface*. Cambridge: University Press.
- Carroll, J. M., Mack R. L. e Kellogg, W. A. (1988). Interface Metaphors and User Interface Design. In M. Helander (a cura di) *Handbook of Human-Computer Interaction*, Amsterdam: North-Holland, pp. 67-105.

- Cassell, J. e Prevost, S. (1996) Distribution of Semantic Features Across Speech & Gesture by Humans and Machine. *Proceedings of the Workshop on the Integration of Gesture in Language and Speech*, October 7-8, Newark and Wilmington Delaware, allegato.
- Catarci, T., Costabile, M. F. e Levialdi, S. a cura di (1992). *Advanced Visual Interface. Proceedings of the International Workshop AVI'92*, Roma, maggio 27-29.
- Chignell, M. H. e Hanckock, P. A. (1988). Intelligent Interface Design. In M. Helander (A cura di) *Handbook of Human-Computer Interaction*, Amsterdam: North-Holland, pp. 969-996
- Cogin, J. P., Diehl, V. A. e Norman, K. L. (1988). Development of an Instrument Measuring User Satisfaction of the Human-Computer Interface. *Proceedings ACM CHI'88 Conference*, pp. 213-218.
- Cohen, B. A. e Waugh, G. W. (1989). Assessing Computer Anxiety. *Psychological Reports*, 65, pp. 735-738.
- Cohen, P. R. e Sullivan, J. W. (1989). Synergistic Use of Direct Manipulation and Natural Language. *Proceedings ACM CHI'89 Conference*, May, pp. 227-233.
- Cole, R. et al., (1995) .The Challenge of Spoken Language Systems: Research directions for the Nineties. *IEEE Transactions on Speech and Audio Processing*, 3, 1, pp. 1-21.
- Coniglio, I. e Paci, A. M. (1994). Usabilità di prodotti: l'esperienza Olivetti. Comunicazione presentata al seminario su *Usabilità di prodotti, servizi e sistemi: centri di ricerca e laboratori italiani d'eccellenza*, Fondazione Ugo Bordoni, Roma, 11 luglio.
- Coover, D. e Delcourt, M. A. B. (1992). Construct and Criterion-Related Validity of the Adult-Attitudes Toward Computers Survey for a Sample of Professional Nurses. *Educational and Psychological Measurement*, 52, pp. 653-661.
- Coutaz, J., Nigay, L. e Salber, D. (1993). Taxonomic Issues for Multimodal and Multimedia Interactive Systems. *Proceedings of ERCIM'93 workshop on Multimedia Multimodal User Interfaces*, Nancy, October.
- Craik, K. J. (1943). *The Nature of Explanation*. New York: Cambridge University Press.
- Dambrot, F. H., Watkins-Malek, A., Silling, M., Marshall, R.S. e Garver, J.A. (1985). Correlates of Sex Differences in Attitude Toward and Involvement with Computers. *Journal of Vocational Behaviour*, 27, pp. 71-86.
- De Angeli, A. (1991). Dillo a Maia: un esperimento di interazione uomo-macchina. Tesi di laurea non pubblicata. Università degli Studi di Trieste, Facoltà di Lettere e Filosofia.
- De Angeli, A. (1993). *Interfacce centrate sull'utente*. Technical Report del Laboratorio Tecnologie Cognitive 93-05.
- De Angeli, A. (1994). *OFFICE 241: alcune misure di usabilità*. Technical Report del Laboratorio Tecnologie Cognitive 94-03.
- De Angeli, A. (1995). Donne nelle reti elettroniche. *Sistemi Intelligenti*, 3, pp. 373-391.
- De Angeli, A., Fum, D. e Gerbino, W. (1995). *Valutazione di sistemi di comunicazione VMS RDS-TMC in ambiente autostradale*. Technical Report del Laboratorio Tecnologie Cognitive, 95-05.
- De Angeli, A., Fum, D., Gerbino, W., Pani, B., Parisatto, B. e Tognoni, G.. (1996). Driver Reaction and Behavior in a Motorway Environment Equipped with an Integrated VMS and On-Board RDS-TMC Information System. In *Proceedings of the*

Third Annual World Congress on Intelligent Transport Systems, Orlando, Florida, ottobre 14-18, disponibile su CD-ROM.

De Angeli, A., Petrelli, D. & Gerbino, W. (1996). Interface Features Affecting Deixis Production: A Simulation Study. In L. Messing (a cura di) *Proceedings of The Workshop on the Integration of Gesture in Language and Speech*, Newark, Delaware and Wilmington, Delaware, October 7-8, pp. 195-204.

De Leo, D., Villa, A., Magni, G., Andreatta, A. e Gagliardi, A. (1986). Presentazione della versione italiana e contributo alla taratura del Bem Sex-Role Inventory. *Bollettino di Psicologia Applicata*, 175, pp. 21-28.

Delin, P. S. (1968). Success in recall as a function of success in implementation of mnemonic instructions. *Psychonomic Science*, 12, pp. 153-154.

Devoto, G. e Oli, G. C. (1990). *Il dizionario della Lingua Italiana*. Firenze: Le Monnier.

Doll, W. J. E Torkzadeh, G. (1988). The measurement of end-user computing satisfaction. *MIS Quarterly*, 12, 259-274.

Dukes R. L., Discenza R. e Couger J. D. (1989). Convergent validity of four computer anxiety scales. *Educational and Psychological Measurement*, 49, pp., 195-202

Egan, D.E, Remde, J. R., Landauer, T. K., Lochbaum, C. C. e Gomez, L. M. (1989). Behavioral evaluation and analysis of a hypertext browser. *Proceedings CHI'89 Human factors in Computing Systems*, pp. 205-210, New York: ACM.

Ericsson, K. A. e Simon, H. A. (1984). *Protocol Analysis: verbal Reports as Data*. Cambridge, MA: MIT Press.

Fetler, M. (1985). Sex Differences on the California Statewide Assesment of Computer Litteracy. *Sex Roles*, 13, pp. 181-191.

Frantzve, J. L., Lips, H. M., Castor, R.R., Wilson, K.A., Kraus, M. e Foster, H. (1992). Attitudes toward computers: effects of gender and power perceptions. Paper presentato al *Twenty-Fifth International Congress of Psychology*, Brussels, Belgium.

Fraser, N. M. e Gilbert, G. N. (1991). Simulating speech system. *Computer, Speech and Languages*, 5, pp. 81-99.

Gaver, B. (1993) "Class you are not making enough noise!" The case for sound-effects in educational software. *Proceedings of Basics of Man-Machine Communication for the Design of Educational System*, Eindhoven, The Netherlands, Agosto 16-26, vol 1, pp.109-120.

Gaver, W. W. (1986). Auditory Icons: Using Sound in Computer Interfaces. *Human-Computer Interaction*, 2, pag 167-177.

Gentner, D. e Nielson, J. (1996). The Anti-Mac Interface. *Communication of the ACM*, 39(8), pp. 70-82.

Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston: Houghton Mifflin.

Gilroy, F. D. e Desay, H. B. (1986). Computer Anxiety: sex age and race. *International Journal of Man-Machine Studies*, 25, pp. 711-71.

Good, M., Spine, T. M., Whiteside, J. e George, P. (1986). The iterative design of a new text editor. *Proceedings ACM CHI'86 Conference*, Boston, April, pp. 241-246.

Gould, J. D, Boies, S. J., Levy, S., Richards, J. T. e Schoonard, J. (1987). The 1984 Olympic Message System: a test of behavioural principle of system design. *Communication of the ACM*, 30(9), pp. 758-769.

- Gould, J. D. (1988). How to design usable system. In M. Helander (a cura di) *Handbook of Human-Computer Interaction*, Amsterdam: North-Holland, pp. 757-789.
- Gould, J. D. e Lewis, C. H. (1985). Designing for usability-key principles and what designer think. *Communication of the ACM*, 28, 300-311.
- Green, T. R. G., Davies, S. P. e Gilmore, D. J. (1996). Delivering cognitive psychology to HCI: the problems of common language and of knowledge transfer. *Interacting with Computers: Journal of Human-Computer Interaction*, 8(1), pp. 89-111.
- Grice, H.P. (1975) Logic and Conversation. In P. Cole e J. L. Morgan (A cura di.), *Syntax and Semantics*, New York: Academic Press.
- Grundin, J (1993). Interface an Evolving Concept. *Communication of the ACM*, 36(4), pp. 112-119.
- Gullberg, M. (1996). Deictic Gesture and Strategy in Second Language Narrative. *Proceedings of the Workshop on the Integration of Gesture in Language and Speech*, October 7-8, Newark and Wilmington Delaware, pp. 155-174.
- Guyomard, M. e Siroux, J. (1988). Experimentation in the specification of oral dialogue. *NATO-ASI Recent Advances in Speech Understandin and dialogue Systems*, F46, pp. 497-501.
- Harker, S. (1995). The development of ergonomics standards for software. *Applied Ergonomics*, 26(4), pp. 275-279.
- Harrison, A. W. e Rainer, R.K., jr. (1992) An examination of factor structure and concurrent validity for the Computer Attitude Scale, the Computer Anxiety Rating Scale, and the Computer Self-Efficacy Scale. *Educational and Psychological Measurement*, 52, pp. 735-745.
- Hauptmann, A. G. e Rudnick, A. I. (1988). Talking to computers: an empirical investigation. *International Journal of Human-Computer Studies*, 28, pp.583-684.
- Hawkins, J. (1985). Computers and girls: rethinking the issues. *Sex Roles*, 13, pp. 165-180.
- Hess, R. D. e Miura, I. T. (1985), Gender differences in Enrollment in Computer Camps and Classes. *Sex Roles*, 13, pp. 193-202.
- Hoiem, D. E. & Sullivan , K. D. (1994). Designing and using integrated data collection and analysis tools: challenges and considerations. *Behaviour 6 Information Technology*, 13, 1-2, 160-170.
- Hutchins, E. L., Hollan, J. D. e Norman, D. A. (1986). Direct Manipulation Interfaces. *Human-Computer Interaction*, 1, pp. 311-338.
- Hutchins, E. L., Hollan, J. D. e Norman, D. A. (1986). Direct Manipulation Interfaces. In D. A. Norman e S. W. Draper (A cura di) *User Centered System Design: New Perspectives on Human-Computer Interaction*, Hillsdale, New Jersey: Lawrence Erlbaum Associates, pp. 88-124.
- Ingwersen, P. (1992). *Information Retrieval Interaction*. London, UK: Taylor Graham.
- Ingwersen, P. (1994). Evaluation of Interactive IR Systems. *Dispensa di seminario-Dipartimento di Informatica, Università di Udine*.
- ISO CD 9241-11.3. (1993) Part 11: Guidance on specifying and measuring usability. In *ISO 9241 Ergonomic requirements for office work with visual display terminals (VDTs)*, pp. 1-88.
- Ives, B. e Olson, M. H. (1984). User involvement and Mis success: a review of research. *Management Science*, 30, pp. 586-630.

- Ives, B., Olson, M. H. e Baroudi, J. J. (1983). The Measurement of User Information Satisfaction. *Communication of the ACM*, 26(10), pp. 785-793.
- Jacob, R. J. K. (1993) New human-computer interaction techniques. In *Proceedings of Basics of Man-Machine Communication for the Design of Educational System*, Eindhoven, The Netherlands, Agosto 16-26, vol 1, pp. 135-142..
- Job, R. e Tabossi, P. (1984). Umana o Artificiale, purchè sia Intelligenza. *Storia e Critica della Psicologia*, 1, pp. 69-89.
- Johnsson, A. e Dahlback, N. (1988). Talking to a computer is not like talking to your best friend. *Proceedings of scandinavian Conference of Artificial Intelligence*, pp. 53-68.
- Kantrowitz, B. (1994). men, women & computers. *Newsweek*, 16 maggio
- Kaptelinin, V. (1993). Activity theory: its implication for man machine communication. In *Proceedings of Basics of Man-Machine Communication for the Design of Educational System*, Eindhoven, The Netherlands, Agosto 16-26, vol 2, pp. 359-367.
- Karat, J. (1988). Software Evaluation Methodologies. In M. Helander (a cura di) *Handbook of Human-Computer Interaction*, Amsterdam: North-Holland, pp. 891-928.
- Kennedy, A., Wilkes, A., Elder, L. e Murray, W. S. (1988). Dialogue with machines. *Cognition*, 30(1), pp.37-72.
- Kennedy, S. (1989). Using Video in the BNR Usability Lab. *ACM SIGCHI Bulletin*, 21, 2, pp. 92-95.
- Kernan, M. C. e Howard, G. S. (1992). Computer Anxiety and Computer Attitudes: An Investigation construct and predictive validity issues. *Educational and Psychological Measuremen*, 50, pp. 681-689
- Kieras, D. E. (1988). Towards a Practical GOMS Model Methodology for User Interface Design. In M. Helander (Ed), *Handbook of Human-Computer Interaction*. Amsterdam: Elsevier Science Publishers, pp. 135-157.
- Landauer, T. K. (1987) Relations between Cognitive Psychology and Computer System Design. In J. M. Carrol, (a cura di) *Interfacing Thought: Cognitive Aspects of Human-Computer Interaction*, Cambridge, MA: MIT Press, pp. 1-25.
- Landauer, T. K. (1988). Research Methods in Human-Computer Interaction. In M. Helander (a cura di) *Handbook of Human-Computer Interaction*, Amsterdam: North-Holland, pp. 905-928.
- Landauer, T. K. e Galotti, K. M. (1984). What makes a difference when? Comments on Grundin and Barnard. *Human Factors*, 26, 423-429.
- Laurel, B. (1991). Interface Agents: metaphors with Character. In B. Laurel (a cura di) *The Art of Human-Computer Interface Design*, pp. 355-365.
- Lea, M. e Spears, R. (1991). Computer-mediated communication, de-individuation, and group decision-making. *International Journal of Man-Machine Studies*, 34 (Special issue on Computer-Supported Cooperative Work and Groupware), pp. 283-301.
- Lea, M. e Spears, R. (1992), Paralanguage and social perception in computer-mediated communication. *Journal of Organizational Computing*, 2, pp. 321-342.
- Lewis, C. H. (1990). A Research Agenda for the Nineties in Human-Computer Interaction. *Human-Computer Interaction*, 5, pp. 125-143.
- Lewis, C. H. e Norman, D. A. (1986). Designing for errors. In D. A. Norman e S. W. Draper *User Centered System Design: New Perspectives on Human-Computer Interaction*, Hillsdale, New Jersey: Lawrence Erlbaum Associates, pp. 411-432.

- Lewis, J. R. (1994). Sample Sizes for Usability Studies: Additional Correlations. *Human Factors*, 36(2), pp. 368-378.
- Licklider, J. C. R. (1960). Man-computer symbiosis. *IRE Transactions on Human Factors in Electronics*, 1(1), pp. 4-11.
- Loyd, B. H. and Gressard, C. (1984). The Reliability and Factorial Validity of Assessment of Computer Attitude Scales. *Educational and Psychological Measurement*, 44, pp. 501-505.
- Loyd, B. H. e Loyd, D. E. (1985). The Reliability and Validity of an Instrument for the Assessment of Computer Attitudes. *Educational and Psychological Measurement*, 45, pp. 903-908.
- Luccio, R. (1982). La psicologia cognitivista. In P. Legrenzia (A cura di) *Storia della Psicologia*, Bologna: Il Mulino, pp. 227-250, nuova edizione.
- Mackinlay J., Card, S.K. e Robertson, G.G. (1990). A Semantic Analysis of the Design Space of Input Devices. *Human-Computer Interaction*, 5(2-3), pp. 145-190.
- Mantovani G. (1993). Comunicazione mediata dal computer, processi di gruppo e cooperazione. *Sistemi Intelligenti*, 10(2), pp. 153-172.
- Marcoulides G. A. (1989). Measuring Computer Anxiety: The Computer Anxiety Scale. In *Educational and Psychological Measurement*, 49, pp. 733-739.
- Marcus, A. (1993). Human Communication Issues in Advanced UIS. *Communication of the ACM*, 4, pp. 101-109
- Marquié, B., Thon, B. e Baracat, B. (1994). Age influence on attitude of office workers faced with new computerized technologies. *Applied Ergonomics*, 25(3), pp. 130-142.
- Matarazzo, G. (1994). I risultati della prima indagine sui laboratori di usabilità in Italia. Comunicazione presentata al seminario su *Usabilità di prodotti, servizi e sistemi: centri di ricerca e laboratori italiani d'eccellenza*, Fondazione Ugo Bordoni, Roma, 11 luglio 1994
- Matherson, K. (1992). Women and computer technology. In M. Lea (a cura di). *Contexts of Computer-Mediated Communication*. London: Harvester Wheatsheaf, capitolo 3.
- Maybury, M.T., A cura di, (1993). *Intelligent Multimedia Interfaces*. Cambridge, Mass: MIT Press
- Mayer, R. E. e Anderson, R. B. (1991). Animation Need Narrations: An Experimental Test of a Dual-Coding Hypothesis. *Journal of Educational Psychology*, 83(4), pp. 484-490.
- Mayer, R. E. e Anderson, R. B. (1992). The Instructive Animation: Helping Students Build Connections Between Words and Pictures in Multimedia learning. *Journal of Educational Psychology*, 84(4), pp. 444-452.
- Mayer, R. E. e Sims, V. K. (1994). For Whom Is a Picture Worth a Thousand Words? Extensions of a Dual-coding Theory of multimedia learning. *Journal of Educational Psychology*, 86, 3, pp. 389-401.
- McNeill, D. (1992). *Hand and Mind: What Gestures Reveal about Thought*. Chicago: University of Chicago Press.
- Miller, G. A., Galanter, E. e Pribram, K. H. (1960). *Plans and structure of behavior*. New York: Holt, Rinhart & Winston; trad. it. *Piani e strutture del comportamento*, Milano: Angeli, 1973.
- Muckler, F. A. (1992). Selecting performance Measures : "Objective" versus "Subjective" Measurement. *Human Factors*, 34(4), 441-155.

- Myers, B. A. (1996). A Quick History of Human Computer Interaction. Draft reso disponibile dall'autore tramite posta elettronica.
- Naas, C., Steuer, J. e Tauber, E. R. (1994). Computers are Social Actors. *Proceedings ACM CHI'94 Conference*, Boston, Massachusetts, April, pp. 72-78.
- Naas, C., Steuer, J., Herinksen, L. e Dryier, D. C. (1994). Machines, social attributions and ethopeia: performance assessments of computers subsequent to "self" or "other" evaluations. *International Journal of Human-Computer Studies*, **40**(3), pp. 403-575.
- Neisser, U. (1967). *Cognitive Psychology*. New York: Appleton, trad. it. *Psicologia Cognitivista*, Firenze: Martello-Giunti, 1976.
- Neisser, U. (1976). *Cognition and Reality*. San Francisco: Freeman & Co.
- Newell, A e Card, S. K. (1985). The prospects for Psychological Science in Human-Computer Interaction. *Human-Computer Interaction*, **1**, pp. 209-242.
- Newell, A e Card, S. K. (1986). Straightening Out Softening Up: Response to Carroll and Campbell Hard Science. *Human-Computer Interaction*, **2**, pp. 251-267.
- Newell, A. e Simon, H. A. (1972). *Human Problem Solving*. Englewood Cliffs: Prentice Hall.
- Nielsen, J. (1993). *Usability Engineering*. San Diego: Academic Press.
- Nielsen, J. (1994) Estimating the number of subjects needed for a thinking aloud test. *International Journal of Human-Computer Interaction*, **41**, pp. 385-397.
- Nielsen, J. (1994). Usability Laboratories. *Behaviour & Information Technology*, **1/2**, pp.3-8.
- Nielsen, J. e Molich, R. (1990). Heuristic evaluation of user interfaces. *Proceedings ACM CHI'90*, Seattle, WA, 1-5 Aprile, pp. 249-256.
- Norman, D. A. (1986). Cognitive Engineering. In D. A. Norman e S. W. Draper (A cura di) *User Centered System Design: New Perspectives on Human-Computer Interaction*, Hillsdale, New Jersey: Lawrence Erlbaum Associates, pp. 31-61.
- Norman, D. A. (1987). Cognitive Science-Cognitive engineering. In J. M. Carroll (a cura di). *Interfacing thought: Cognitive Aspect of human-computer interaction*. Cambridge MA: Brandford, MIT Press.
- Norman, D. A. (1988). *The Psychology of Everyday Things*. New York: Basic Books.
- Norman, D. A. (1991). Cognitive Artifacts. In J. M. Carroll (a cura di) *Designing Interaction: psychology at the Human-Computer Interface*. Cambridge: University Press. Tr. it. Artefatti cognitivi, *Sistemi intelligenti*, **3**, pp. 453-476, 1991.
- Norman, D. A. e Draper, S. W. (1986). *User Centered System Design: New Perspectives on Human-Computer Interaction*. Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- Nuttin, J. (1980). *Teoria della motivazione Umana*. Roma: Armando Armando.
- O'Malley, C. E., Draper, S. W. e Riley, M. S. (1984). Constructive interaction: A method for studying human-computer-human interaction. *Proceeding IFIP INTERACT'84 First International Conference on Human- Computer Interaction*, London, U.K., Settembre 4-7, pp. 269-274.
- Ole Bersen, N. (1993). *Taxonomy of HCI Systems: State of the Art*. GRACE - EXPRIT Basic Research, Deliverable 2.1, Agosto.

- Olson, J. R. e Olson, G. M. (1990). The Growth of cognitive Modelling in Human-Computer Interaction Since GOMS. *Human-Computer Interaction*, 2, pp. 221-265.
- Oviatt, S. (1995). Predicting spoken disfluencies during human-computer interaction. *Computer Speech and Language*, 9, pp. 19-35.
- Oviatt, S. (1996). Multimodal Interfaces for Dynamic interactive Maps. *CHI'96 Conference Proceedings*. Denver: Colorado.
- Oviatt, S. e Olsen, E. (1994). Integration themes in multimodal human-computer interaction. *Proceedings of the International Conference on Spoken Language Processing*. Yokohama, Japan. pp. 551-554.
- Oviatt, S. L., (1994). Interface Techniques for Minimizing Disfluent Input to Spoken Language Systems. *Proceedings of Human Factors in Computing Systems*, Boston, Massachusetts, USA, April, pp. 24-28.
- Oviatt, S. L., Cohen, P. R. e Wang, M., (1994) Toward interface design for human language technology: Modality and structure as determinants of linguistic complexity, *Speech Communication*, 15, pp. 283-300.
- Oviatt, S. L., Cohen, P. R., Fong, M. e Frank, M. (1992). A rapid semi-automatic simulation technique for investigating interactive speech and handwriting. In J. Ohala (a cura di), *Proceedings of the 1992 International Conference on Spoken Language processing*, vol. 2, Alberta, pp. 1351 -1354.
- Oviatt, S. L., Cohen, P. R., Fong, M. e Frank, M. (1992). A rapid semi- automatic simulation technique for investigating interactive speech and handwriting. In J. Ohala, (ed) *Proceedings of the 1992 International Conference on Spoken Language Processing*, vol. 2, University of Alberta, pp. 1351-1354.
- Oviatt, S., De Angeli, A. e Kuhn, K. (1997). Integration and Synchronization of Input Modes during multimodal Human-Computer Interaction. Articolo accettato a ACM CHI'97 Conference, Atlanta, Marzo.
- Paap, K. R. e Roske-Hofstrand, R. J. (1988). Design of Menus. In M. Helander (A cura di) *Handbook of Human-Computer Interaction*, Amsterdam: North-Holland, pp. 205-235..
- Parasumaran, S. e Igarria, M. (1990). An examination of gender differences in the determinants of computer anxiety and attitudes toward microcomputers among managers. *International Journal of Man-Machine Studies*, 32, pp. 327-340.
- Parisi, D. (1989). *Intervista sulle reti neurali*. Bologna: Il Mulino.
- Polson, P. G. e Kieras, D. E. (1985). A quantitative model of the learning and performance of text editing knowledge. *Proceedings ACM CHI'85 Conference*, pp. 207-212.
- Popovich, P. M., Hide, K. R., Zakrajsek D. e Blumer, C. (1987). The development of the Attitudes toward Computer Usage Scale. *Educational and Psychological Measurement*, 47, pp. 261-269.
- Preece, J., Rogers, Y., Sharp, H., Benyon, D., Holland, S. e Carey, T. (1994). *Human-Computer Interaction*. Reading, MA: Addison-Wesley.
- Rasmussen, J. (1983). Skills, Rules, Knowledge: Signals, Signs and Symbols and other distinctions in human performance models. *IEEE Transactions: Systems, Man & Cybernetics*, SMC 13, pp. 257-267.
- Rasmussen, J. (1986). *Information processing and Human-Machine Interaction: An Approach to Cognitive Engineering*. New York: North-Holland.

- Re, A. (1995). *Ergonomia per psicologi. Lavoro Cognitivo e Nuove Tecnologie*. Milano: Raffaello Cortina Editore.
- Reason, J. (1990). *Human Error*. New York: Cambridge University Press.
- Reithinger, N. (1987). Generating referring expressions and pointing gestures. In G. Kempen (a cura di) *Natural Language Generation*. Dordrecht: Kluwer, pp. 71-81.
- Rieber, L. P. (1990). Using Computer Animated Graphics in Science Instruction With Children. *Journal of Educational Psychology*, **82**(1), pp. 135-140.
- Rizzo, A., Bagnara, S. e Visciola, M. (1987). Human Error Detection Processes. *International Journal of Man-Machine Studies*, **27**.
- Robertson, S. e e Hancock-Beaulieu, M. M. (1992). On the evaluation of IR Systems. *Information Processing & Management*, **28** (4), pp. 457-466.
- Rosenberg, R. (1991). Debunking computer literacy. *Technology Review*, **94**(1), pp. 58-65.
- Rumiati, R. (1984). Il pensiero e la soluzione dei problemi. In R. Job e R. Rumiati *Linguaggio e Pensiero*. Bologna: Il Mulino, capitolo 4.
- Sacks, H. A., Schegloff, E. A. e Jefferson, G. (1974). A simplest systematic for the Organization of Turn taking in Conversation. *Language*, **50**, pp. 696-735.
- Salber, D. e Coutaz, J. (1995). Applying the Wizard of Oz Technique to the study of Multimodal System. *International Journal of Human-Computer Interaction*, **4**, 219-230.
- Salton, G. (1992). The State of Retrieval System Evaluation. *Information Processing & Management*, **28** (4), pp. 441-449.
- Sanders, M. S. e McCormick, E. J. (1992). *Human Factors in Engineering and Design*. Seventh Edition. New York: McGraw-Hill, Inc.
- Schmauks, D. (1987). Natural and simulated pointing. *Proceedings of the 3rd European ACL conference*, Copenhagen, Denmark, pp. 179-185.
- Schmauks, D. e Reithinger, N. (1988). Generating Multimodal Output-Conditions, Advantages and Problems. *Proceedings of the International Conference on Computational linguistics*, Budapest, Hungary, pp. 584-588.
- Schmauks, D. e Wille M. (1991) Integration of Communicative Hand Movements into Human-Computer-Interaction. *Computers and the Humanities*, **25**, pp. 129-140.
- Schooler, J. W. e Englsteler-Schooler T. Y. (1990). Verbal overshadowing of visual memories: Some things are better left unsaid. *Cognitive Psychology*, **22**, pp. 36-71.
- Schooler, J. W., Ohlsson, S. e Brooks, K. (1993). Thoughts bejound words: When language overshadows insight. *Journal of Experimental Psychology: General*, **122**, pp. 166-183.
- Sensales, G. e Greenfield, P. (1991), Computer, Scienza e Tecnologia: un confronto trans-culturale fra gli atteggiamenti di studenti italiani e statunitensi. *Giornale Italiano di Psicologia*, **1** a. XVIII., pp. 45-57.
- Shneiderman, B. (1983). Direct manipulation: A step beyond programming languages. *IEEE Computer*, **16**(8), pp. 57-69.
- Shneiderman, B. (1992). *Designing the User Interface-strategies for Effective Human-Computer Interaction*. Reading, MA: Addison-Wesley, seconda edizione.
- Simon, H. A. (1969). *The Sciences of the Artificial*. Cambridge, Mass.: MIT Press.

Simon, H. A. (1975). The functional equivalence of problem-solving skills. *Cognitive Psychology*, **7**, pp.268-288.

Spears R. e Lea, M. (1994),. *Panacea or Panopticon? The hidden Power in Computer-Mediated Communication. Communication Research*, **21**(4), pp.. 427-459.

Spears, R. e Lea, M. (1992). Social influence and the influence of the 'social' in computer-mediated communication. In M. Lea (A cura di) *Contexts of Computer-Mediated Communication*, London: Harvester-Wheatsheaf.

Spears, R., Lea, M. e Lee, S. (1990). De-individuation and group polarization in computer-mediated communication. *British Journal of Social Psychology*, **29**, pag 121-134.

Stewart, T. (1995). Visual displays, controls and environmental requirements. *Applied Ergonomics*, **26**(4), pp. 271-274.

Stock, O. (1987a). Prefazione. In O. Stock (A cura di) *L'interazione uomo-computer*, Roma: Il Pensiero Scientifico Editore, pp. 1-3.

Stock, O. (1987b). Interazione uomo-macchina intelligente. In O. Stock (a cura di) *L'interazione uomo-computer*, Roma: Il Pensiero Scientifico Editore, pp. 3-17

Stock, O. (1995). A third modality? *Artificial Intelligence Review*, **9**, pp. 129-146

Stock, O., Strapparava, C. e Zancanaro, M. (1995). Explorations in a Natural Language Multimodal Information Access Environment. *Proceedings of IJCAI 95 Workshop on Intelligent Multimodal Interfaces*.

Su, L.S. (1992). Evaluation measures for interactive information retrieval. *Information Processing & Management*, **28**(4), pp. 503-516.

Szaina, B (1994), An Investigation of the Predictive Validity of Computer Anxiety and Computer Aptitude. *Educational and Psychological Measurement*, **5**, pp. 926-936.

Tabossi, P. (1988). *Intelligenza naturale e intelligenza artificiale*. Introduzione alla scienza cognitiva. Bologna: Il Mulino.

Tague-Sutcliffe, J. (1992). The pragmatics of Information retrieval experimentation, revisa cura di *Information Processing & Management*, **28**(4), pp. 467-490.

Thimbleby, J. (1990). *User Interface Design*. ACM Press: New York.

Turner, J. C., Hogg, M. A, Oakes, P. J., Reicher, S. e Wetherell, M. (1987). *Rediscovering the Social Group. A Self-Categorization Theory*. Oxford: Basil Blackwell.

van nes, F. L. e van Itegem, J. P. (1990). Hidden functionality: how an advanced car radio is really used. *IPO Annual Progress Report*, **25**, pp. 101-121.

van Vliet, P. J., Kletke, M. G. e Chakraborty, G. (1994). The measurement of computer literacy: a comparison of self-appraisal and objective tests. *International Journal of Human- Computer Studies*, **40**, pp. 835-875.

Varalda, G., Bray, R., Morra, G.L., Sidoti, C., Gerbino, W., De Angeli, A. e Ferrante, D. (1994). Evaluation of Application 2: Collision Warning and Medium Range Warning Simulator Experiment. In *Evaluation of Results*, Deliverable n° 10, Type: Reserved, Drive II Project V2006 EMMIS, dicembre, capitolo 4.

Virzi, R. A. (1992). Refining the Test Phase of Usability Evaluation: How Many Subjects Is Enough?. *Human Factors*, **34**(4), pp. 457-468.

Wærn, Y. (1989). *Cognitive aspects of computer supported tasks*. New York: John Wiley & Sons.

Wagman, M. (1983). A factor analytic study of the psychological implications of the computer for the individual and society. *Behavior Research Methods & Instruments*, **15**(4), pp. 413-419.

Wahlster, W. (1992). Automatic design of Multimodal Presentations. In Catarci, T., Costabile, M. F. e Levialdi, S. (A cura di) *Advanced Visual Interfaces Proceedings of the International Workshop AVI'92*.

Ware, M. C. e Stuck, M. F. (1985). Sex-Role Messages vis-à-vis Microcomputer Use: A Look at the -Pictures. *Sex Roles*, **13**, pp. 205-228.

Whiteside, J., Bennet, J. e Holtzblatt K. (1988). Usability Engineering: Our Experience and Evolution. In M. Helander, (a cura di) *Handbook of Human-Computer Interaction*, North Holland: Amsterdam, pp. 791-817.

Winter, H. (1986). Artificial Intelligence in Man-Machine Systems and Automation. *Proceedings of an International Seminar Organized by DFVLR*, Bonn, pp. 1-23.

Woods, D. D.e Roth, E. M. (1988). Cognitive Systems Engineering. In M. Helander (a cura di) *Handbook of Human-Computer Interaction*, Amsterdam: North-Holland, pp. 3-43.

Wright, P. (1989). The need for a theories of NOT reading: some psychological aspects of the human-computer interface. In B.A.G Elsendoorn e H. Bouma (a cura di) *Working Models of Human Perception*, pp. 319-340, London: Academic Press.

Zancanaro, M., Stock, O. e Strapparava, C. (in stampa) Multimodal interaction for Information Access: Exploiting Cohesion. *Computational Intelligence*.

Zancanaro, M., Stock, O.e Strapparava, C. (1993). Dialogue Cohesion Sharing and Adjusting in An Enhanced Multimodal Environment. In *Proceedings of 13th International Joint Conference on Artificial Intelligence*, Chamberey, France, pp. 1230-1236.

Zancanaro, M., Stock, O.e Strapparava, C. (1994). To the Rescue of the Lost Explorer: Steps Toward Overcoming the Problems in Information Access Dialogues. *Proceedings of AAAI94 Spring Symposium on Intelligent Multi-Media Multi-Modal Systems*.