

UNIVERSITÀ DEGLI STUDI DI SALERNO



Interazione Uomo Macchina e Usabilità del Software

MedMinder

Assignment 3

Usability Testing

Francesca Festa -

Emanuele Gargiulo - 0522500718

Indice

INDICE	2
INTRODUZIONE	4
Definizione ISO dell'Usabilità	4
Valutazione Empirica dell'usabilità	5
IMPOSTAZIONE DEL TESTING PER MEDMINDER	6
Metodologia da Utilizzare	6
Euristiche di Nielsen	7
Requisiti di usabilità da testare	8
Task da testare	8
SCELTA DEI PARTECIPANTI	8
Questionario di background	9
Questionario di Reazione	10
IPOTESI DA TESTARE	12
Metodologia	12
Esperimento numero 1 – verificare come le conoscenze tecnologiche possano influenzare le performance dell'utente nello svolgimento dei task.	12
ISTRUZIONI SUL TESTING E SUI TASK	13
Istruzioni generali sullo studio di usabilità	13
Istruzioni specifiche per i task	13
METODO DI RACCOLTA	14
VALIDITÀ ECOLOGICA DELLO STUDIO	16
METRICHE	17
ANALISI DEI RISULTATI	18
Valutazione Euristica	19
Risultati dei test con gli utenti	21
	2

Successo dei test	21
Tempo impiegato	22
Numero errori	23
Risultati operazioni tramite feedback dell'utente	24
Soddisfazione utenti	25
Invalidazione ipotesi nulla	26
CONCLUSIONI	27
RIFERIMENTI	28

Introduzione

Definizione ISO dell'Usabilità

Secondo la definizione ufficiale dell'ISO 9241-11, l'usabilità è "la misura in cui un prodotto può essere utilizzato da utenti specifici per ottenere determinati obiettivi con efficacia, efficienza e soddisfazione.

- **Efficacia:** indica la capacità di raggiungere l'obiettivo prefissato. Si dice quindi che un prodotto software è efficace se permette di concludere un task che ci si era prefissati; nel caso in cui invece il task non viene completato con successo, l'efficacia si calcola considerando il numero di operazioni che si svolgono per portare a termine il task e la qualità del risultato ottenuto.
- **Efficienza:** valuta l'abilità di portare a termine un compito impiegando le risorse minime indispensabili. Principalmente per l'efficienza vengono considerati due fattori:
 - *tempo di utilizzo:* indica il tempo totale di utilizzo del sistema, per eseguire un task che ci si era prefissati. Un sistema risulta efficiente se il tempo di utilizzo risulta basso.
 - *spazio occupato:* indica la quantità di spazio in memoria occupata dal prodotto software. Minore è lo spazio occupato, maggiore risulta essere l'efficienza.

Per comprendere al meglio la differenza tra efficacia ed efficienza, occorre fare una distinzione. Considerati due sistemi A e B, se entrambi i sistemi portano a termine i task, sono considerati entrambi efficaci; se invece il sistema B porta a termine il compito in un minor tempo e con meno passi, è considerato efficiente rispetto al sistema A.

- **Soddisfazione:** processo volto a rilevare quanto è soddisfatto l'utente rispetto a un prodotto offerto. Il concetto di soddisfazione è correlato alle aspettative dell'utente e alla percezione della qualità del prodotto software offerto, al fine di migliorare il sistema avviando processi di rimodulazione dei servizi sulla base dei feedback raccolti. Il metodo più opportuno per misurare il grado di soddisfazione degli utenti è quello di raccogliere i loro pareri mediante questionari o interviste.

Valutazione Empirica dell'usabilità

La valutazione empirica è lo standard utilizzato per effettuare la valutazione dell'usabilità, poiché permette di capire cosa succede quando il sistema è usato dagli utenti nel mondo reale. Questo metodo può risultare spesso rischioso perché se si attende troppo per far valutare il sistema all'utente, ossia se il sistema è già in una fase avanzata di progettazione, si rischia di dover rifare tutto daccapo.

Esistono diversi modi per portare avanti le valutazioni empiriche. Un primo metodo è quello dei *field studies* in cui si studia il sistema nell'ambito in cui deve essere effettivamente usato, raccogliendo quindi informazioni e classificandole secondo un grado di priorità in base ai problemi riscontrati. Questo procedimento è molto dispendioso in termini di tempo che economici. Per rendere più economico questo processo, si può pensare di far eseguire queste valutazioni agli utenti in laboratorio, assicurandosi che il risultato finale non sia condizionato dall'ambiente in cui si è svolto il test.

Un secondo metodo è quello delle *interviste retrospettive*, in cui si domanda all'utente di ricordare operazioni effettuate col sistema che risultano positive o troppo negative. Una pecca di questa tecnica è che si chiede all'utente di procedere mnemonicamente, il che porta spesso a problemi di inaffidabilità.

Lo scopo quindi del test di usabilità è quello di individuare eventuali criticità e colli di bottiglia dell'interfaccia o delle funzionalità del prodotto software, per poterli poi correggere. Questo tipo di test serve anche a capire come l'utente si muove e ragiona per capire quindi quali sono le eventuali difficoltà.

Per comprendere il ragionamento dietro le azioni che l'utente esegue, gli si chiede di pensare ad alta voce, utilizzando quindi la tecnica del **thinking aloud** che prevede che l'utente spiega quali sono i propri obiettivi, le reazioni alle risposte del sistema e tutti i dubbi che gli vengono in mente durante l'interazione.

L'analisi dei dubbi espressi dall'utente in questa fase può aiutare a scoprire le cause e le conseguenze dei problemi di usabilità riscontrati dall'utente.

Impostazione del testing per MedMinder

Metodologia da Utilizzare

Avendo ricevuto il sistema da testare in una fase di progettazione avanzata, o nella quale il sistema è addirittura già completo, la scelta di utilizzare test empirici per la valutazione è quasi obbligata.

La principale metodologia di valutazione adottata consisterà dunque nel far svolgere agli utenti una selezione di task principali in una situazione che simuli quanto più possibile il reale ambiente d'uso del sistema.

Mentre gli utenti svolgono tali task, i valutatori (in questo caso noi) prenderanno nota di tutte le azioni svolte dagli utenti – eventualmente, anche registrando video dell'interazione tra utente e sistema - al fine di poterle riesaminare per comprendere le motivazioni dietro i principali problemi incontrati nello svolgimento dei task.

Inoltre, agli utenti verrà chiesto di **pensare ad alta voce** (think aloud). In questo modo, oltre alle azioni effettivamente svolte dagli utenti, i valutatori avranno modo di esaminare anche i ragionamenti che hanno portato a tali azioni ed eventuali dubbi che gli utenti si sono posti durante lo svolgimento dei task.

Oltre ai valutatori, con i quali l'utente non potrà interagire durante il test, è opportuno prevedere la presenza di un **facilitatore**, che possa fornire all'utente alcune informazioni o istruzioni di massima durante la prova, aiutando l'utente a superare situazioni in cui potrebbe non riuscire più ad andare avanti con i task.

I risultati ottenuti mediante questa valutazione empirica potranno essere eventualmente integrati con i risultati di una valutazione euristica, cioè verificando la corrispondenza del sistema ad alcune semplici 'regole' individuate da esperti del settore tra cui Nielsen, colui che ha elaborato le euristiche forse più celebri – di seguito riportate.

I problemi d'usabilità nascono quando l'utente non riesce ad intuire le modalità di funzionamento del prodotto e deve necessariamente sforzarsi ad eseguire azioni molto lontane dalle sue aspettative e dalle sue normali modalità di esecuzione. Per comprendere al meglio quali siano gli aspetti positivi di un sistema e quali invece gli aspetti negativi, si possono seguire le 10 euristiche di Nielsen.

1 – Visibilità dello stato del sistema

- Il sistema deve sempre tenere informato l'utente su cosa sta facendo, fornendo un adeguato feedback in un tempo ragionevole.

2 – Corrispondenza tra sistema e mondo reale

- Il sistema deve parlare il linguaggio dell'utente, con parole, frasi e concetti a lui familiari.

3 – Controllo e libertà

- L'utente deve avere il controllo del contenuto informativo e deve potersi muoversi liberamente.

4 – Consistenza e standard

- L'utente deve aspettarsi che le convenzioni del sistema siano valide per tutta l'interfaccia.

5 – Prevenzione dell'errore

- Evitare di porre l'utente in situazione ambigue, critiche e che possono portare all'errore.
- Dare la possibilità di tornare indietro

6 – Riconoscimento anziché ricordo

- Produrre layout semplici e schematici
- Evitare che l'utente riscopra ogni volta l'interfaccia

7 – Flessibilità d'uso

- Offrire delle scorciatoie per i più esperti

8 – Disegno e estetica minimalista

- Dare maggior importanza al contenuto che all'estetica.
- Evitare di accentuare oggetti irrilevanti o raramente necessari (immagini grandi, etc.)
- Evitare che il contenuto informativo della pagina sia messo in secondo piano
- Evitare che l'utente si distraiga o si confonda

9 – Aiuto all'utente

- I messaggi di errore devono essere espressi in linguaggio comprensibile e suggerire soluzioni
- Chiedere conferma per un'azione importante

10- Documentazione

- Anche se il sistema dovrebbe essere usabile senza documentazione è preferibile che essa sia disponibile

Requisiti di usabilità da testare

In accordo con il progettista di MedMinder, i requisiti di usabilità da testare saranno:

- **Efficacia:** L'utente deve portare a termine tutti i task principali con scarse percentuali d'errore.
- **Efficienza:** L'utente deve portare a termine tutti i task principali velocemente.
- **Soddisfazione:** L'utente deve utilizzare il sistema con un sufficiente grado di soddisfazione.

Tali parametri verranno testati per i task principali dell'applicazione.

Task da testare

1. Aggiunta di un farmaco all'applicazione;
2. Aggiunta di una visita all'applicazione;
3. Visualizzazione del riepilogo delle medicine rimanenti;
4. Visualizzazione del riepilogo delle visite mediche;
5. Visualizzazione farmacie aperte in zona;
6. Visualizzazione riepilogo generale.

Scelta dei Partecipanti

La scelta dei partecipanti è cruciale affinché i risultati ottenuti dai test di valutazione siano attendibili.

Sarà utile suddividere gli utenti in due macrocategorie: quelli che hanno una certa familiarità con la tecnologia e quelli che invece non ne hanno.

Tutti gli utenti selezionati dovrebbero avere un potenziale interesse nell'utilizzo di un sistema come MedMinder.

I partecipanti dunque saranno selezionati tra colleghi dell'Università di Salerno - che rappresenteranno, verosimilmente, gli utenti con skill tecnologiche - e membri della famiglia dei valutatori, che rappresenteranno verosimilmente utenti con skill tecnologiche basilari o assenti.

Ad ogni modo, a ciascun utente verrà sottoposto un questionario di background per verificare tra le varie proprie il suo livello di conoscenza con la tecnologia – oltre che il livello di istruzione – al fine di poter analizzare i risultati ottenuti mettendoli potenzialmente in correlazione anche con questi aspetti.

I partecipanti scelti sono in totale 10.

Questionario di background

Come accennato nella sezione precedente, a ciascuno dei partecipanti verrà consegnato un questionario di background da compilare, a fine di raccogliere maggiori informazioni che ci consentiranno poi di individuare eventuali correlazioni tra il background tecnologico/culturale ed i risultati dei test.

Questionario di Background			
Nome	Cognome	Data	
Età:	Sesso:		
Titolo di Studio			
Diploma Superiore	Laurea Triennale	Laurea Magistrale	Ambito: (Scientifico, Umanistico...)
Skill tecnologiche			
Quanto è familiare con l'utilizzo di uno smartphone?			
Pochissimo	Poco	Mediamente	Molto

Questionario di Reazione

In questa fase vengono proposte all'utente delle affermazioni relative all'uso del sistema e secondo la Scala di Likert l'utente dovrà esprimere quanto è d'accordo con tale affermazione in un range che va da "Fortemente in disaccordo" a "Fortemente in accordo".

Un questionario di questo tipo permette di valutare anche aspetti dell'usabilità che potrebbero non risultare da una valutazione puramente numerica come quella basata sulla percentuale di errori o sul tempo di esecuzione dei task.

In aggiunta a tale questionario, verranno poste all'utente delle **domande a risposta aperta** dandogli la possibilità di rispondere ad esse liberamente, per raccogliere ulteriori informazioni.

L'applicazione è semplice da usare				
Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree

Mi sono sentito a mio agio nell'utilizzare l'applicazione				
Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree

Ho trovato incoerenze tra le diverse funzionalità dell'applicazione				
Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree

Mi è stato subito chiaro lo scopo dell'applicazione				
Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree

Userò realmente l'applicazione, o la userei se ne avessi bisogno				
Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree

I messaggi mostrati a schermo sono sempre chiari				
Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree

È facile capire come svolgere ciascuna operazione				
Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree

Sono soddisfatto della mia esperienza d'uso con l'applicazione				
Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree

Trovo che le funzionalità dell'applicazione siano bene integrate				
Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree

Ho avuto necessità di tempo per capire come utilizzare bene l'applicazione				
Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree

Quali sono state le tre caratteristiche più interessanti del sistema? E le tre peggiori?

Cosa si potrebbe migliorare?

Ha altri consigli / feedback che desidererebbe fornire al team di sviluppo?

Ipotesi da Testare

Metodologia

Verranno formulate delle ipotesi riguardo i risultati degli esperimenti, affermando che una variazione nelle variabili indipendenti (*IV*) causerà una variazione delle variabili dipendenti (*DV*).

Per fare ciò, sarà necessario invalidare l'ipotesi nulla – ossia, l'ipotesi per la quale le variabili dipendenti in realtà non vengono modificate in seguito a un cambiamento in quelle dipendenti.

Esperimento numero 1 – verificare come le conoscenze tecnologiche possano influenzare le performance dell'utente nello svolgimento dei task.

Ipotesi nulla H_{01}: non c'è differenza in termini di efficacia tra gli utenti che hanno skill tecnologiche avanzate e skill tecnologiche di base.	
Variabile Indipendente	Variabili Dipendenti
Background Tecnologico degli Utenti	Efficacia
	Efficienza
	Soddisfazione

Dato che il sistema da testare ci è stato fornito a sviluppo già concluso, non è risultato possibile testare ipotesi nulle relative alla scelta di variabili indipendenti relative all'interfaccia (esempio, non è stato possibile testare se l'efficienza varia ingrandendo un determinato bottone).

Data la natura dell'ipotesi nulla individuata, non è in realtà necessario assegnare specifici ordini di esecuzione dei task per i vari utenti, in quanto non si presenta il problema dei **learning bias**.

Istruzioni sul testing e sui task

Istruzioni generali sullo studio di usabilità

‘Nei prossimi 10 minuti, verranno svolti una serie di test relativi all’applicazione MedMinder.

Saranno fornite istruzioni più specifiche per ciascun task, ma tali istruzioni non saranno dettagliate, in quanto uno degli scopi del test è verificare quanto facilmente tale applicazione può essere utilizzata da persone che non hanno conoscenze pregresse del sistema.

Se in un qualunque momento doveste incontrare problemi, seguite semplicemente il vostro istinto, utilizzando le informazioni che avete. Solo se necessario, un facilitatore interverrà per aiutarvi a compiere progressi.

Per favore, ricordatevi di ‘pensare ad alta voce’, dicendo sempre anche quando si sta iniziando lo svolgimento di un certo task e quando lo si è completato.

Ciò è molto importante per comprendere i vostri obiettivi, le vostre aspettative e le vostre reazioni durante l’esecuzione dei task. Se avete domande prima di iniziare, questo è il momento giusto per porle’.

Istruzioni specifiche per i task

Background (per tutti i task): siete una persona affetta da svariate patologie, per le quali dovete assumere diversi medicinali nel corso di una giornata tipo e affrontare molte visite mediche nell’arco dell’anno. Quest’applicazione vuole aiutarvi a tenere sotto controllo le scorte delle vostre medicine e a ricordarvi quando prenderle, così come vuole aiutarvi a ricordare visite mediche e a trovare la farmacia aperta più vicina

Task 1: Aggiunta di un farmaco all’applicazione.

Avete appena acquistato un farmaco, e ora dovete aggiungerlo all’applicazione. Sarà necessario indicare la quantità di farmaco a vostra disposizione e gli orari in cui dovrete assumerlo.

Task 2: Aggiunta di una visita all’applicazione.

Avete appena preso un appuntamento per una visita specialistica. Dovrete aggiungere i dettagli di tale appuntamento all’applicazione, in modo che essa possa ricordarvene.

Task 3: Visualizzazione del riepilogo delle medicine rimanenti.

Non ricordate esattamente se avete abbastanza antibiotici per completare l’intero ciclo di trattamento. Volete controllare sull’applicazione quanti ve ne rimangono.

Task 4: Visualizzazione del riepilogo delle visite mediche.

Non ricordate esattamente quali visite avete questo mese e volete dunque visualizzare tutti i vostri appuntamenti all'interno dell'applicazione.

Task 5: Visualizzazione farmacie aperte in zona.

Vi siete resi conto che vi servirà dell'altro antibiotico in seguito ad un'estrazione dentale. Dovendolo acquistare, cercate la farmacia aperta più vicina utilizzando l'applicazione.

Task 6: Visualizzazione riepilogo generale.

Desiderate dare uno sguardo rapido al prossimo appuntamento medico al quale recarvi e a quale farmaco dovrete assumere più imminente.

Metodo di Raccolta

Prima di iniziare il test vero è proprio, è stato opportuno spiegare a ciascuno dei partecipanti che tipo di applicazione avrebbero utilizzato, e quali sono i suoi scopi.

Bisogna inoltre sottolineare che è normale, in fase di testing, individuare dei problemi: non c'è niente di cui vergognarsi per l'utente nell'ammettere di avere difficoltà, in quanto riscontrarne aiuterà lo sviluppatore a migliorare l'usabilità dell'applicazione.

Per evitare agli utenti lo stress di affrontare il test senza aver mai visto il sistema, essi potranno farvi cinque minuti di pratica precedentemente al test vero e proprio, discutendo con il facilitatore eventuali dubbi generici sulle funzioni o sullo scopo del sistema.

Tuttavia, durante tali minuti di pratica, non sarà consentito fare domande specifiche e relative allo svolgimento dei task al facilitatore ai valutatori che presenzieranno: se fossero questi ultimi a risolvere tutti i dubbi dell'utente, potenziali problemi di usabilità potrebbero non essere individuati e falsare i risultati del test.

Ciascun valutatore avrà a disposizione il seguente template, che utilizzerà per annotare vari aspetti relativi al test in uso.

UTENTE:	TASK:	Ora Inizio:	Ora Fine:
Che schermata sta visualizzando l'utente?			
Che dubbi sta esprimendo al riguardo?			
Le ipotesi che sta formulando su come proseguire, sono corrette?			
Quali difficoltà sta riscontrando?			
Quali ambiguità del sistema hanno rischiato di confondere l'utente, se ve ne sono?			

In aggiunta a tale template, sarà utile registrare in maniera sincrona sia lo schermo del cellulare che il viso e le parole dell'utente. In tal modo, sovrapponendo le due registrazioni, sarà possibile esaminare le sue reazioni a ogni schermata presentata dal sistema, potendo anche ascoltare i ragionamenti dell'utente – dato che gli è stato chiesto di pensare ad alta voce.

Durante lo svolgimento della prova, non sarà possibile alcuna interazione con i valutatori, che dovranno potersi concentrare pienamente sugli appunti da prendere.

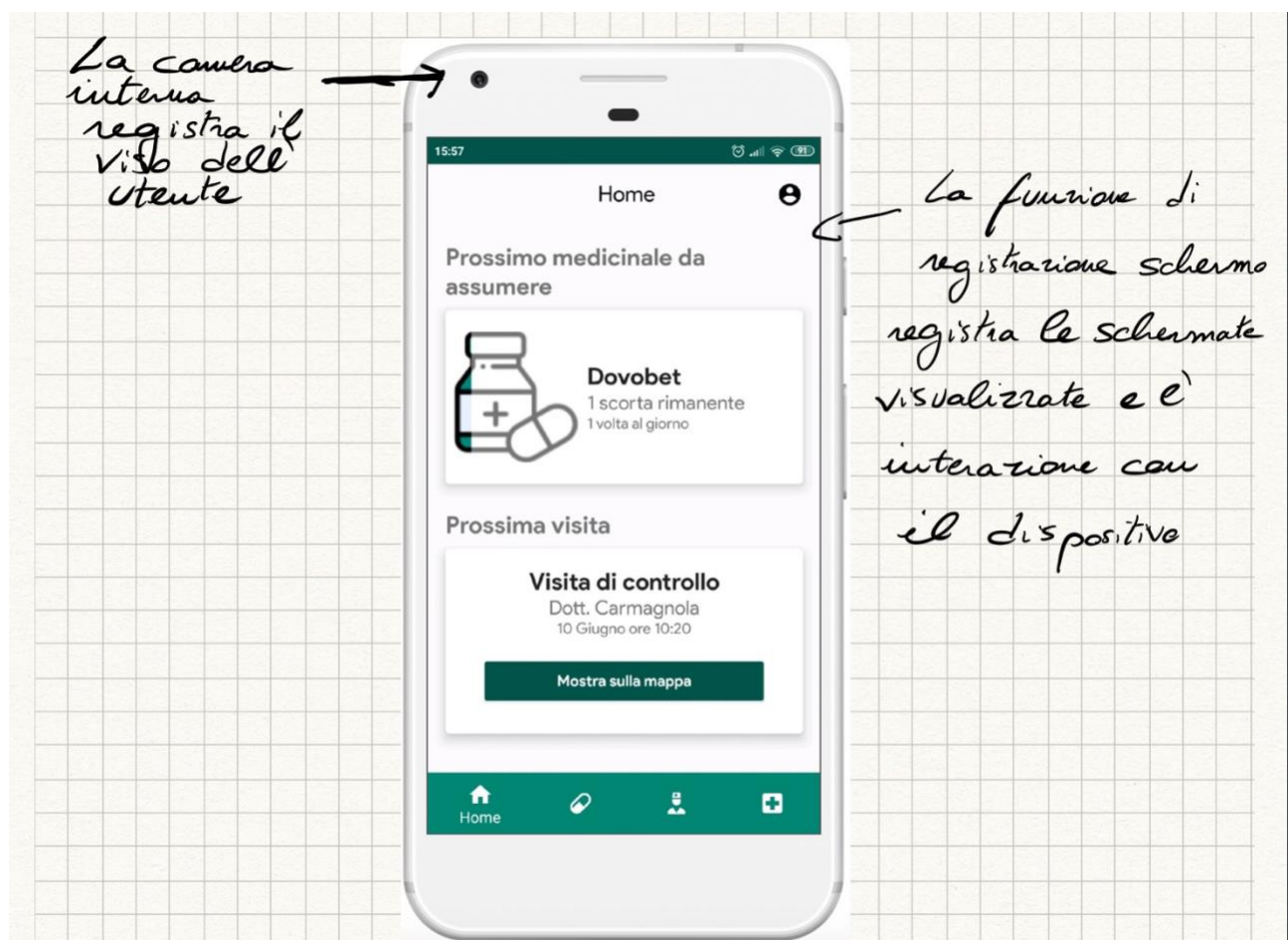
L'interazione con il facilitatore dovrà a sua volta essere limitata al minimo necessario: egli fornirà rassicurazioni e incoraggerà l'utente al **think-aloud**, senza però interferire con lo svolgimento del task.

Conclusa l'interazione con il sistema, alla fine del test ad ogni utente sarà sottoposto il questionario di reazione costruito in precedenza, al fine di ottenere informazioni sul sistema che potrebbero non essere strettamente legate al successo o all'efficienza con la quale l'utente ha svolto i task.

Validità Ecologica dello Studio

Dato che l'applicazione in uso verrà eseguita direttamente su uno smartphone, e che tutti gli smartphone degli ultimi anni dispongono di una fotocamera interna – oltre che della possibilità fornita tramite API di registrare lo schermo – per avere un'ambiente il più possibile simile a quello di utilizzo reale, la registrazione delle schermate e dell'utente potrebbero avvenire semplicemente attivando sia la registrazione schermo del dispositivo che la registrazione del volto mediante fotocamera interna.

Tali schermate andrebbero in seguito sovrapposte per ottenere un video unico e sincrono dell'interazione tra utente e sistema.



Tuttavia, nel caso specifico del test portato avanti, non si disponeva della versione completamente implementata dell'applicazione né della possibilità di integrare in essa le funzionalità di registrazione mediante schermo e fotocamera.

Disponendo del prototipo in Marvel, è stato deciso di avviare quest'ultimo direttamente su smartphone – per ottenere un'esperienza d'uso il più realistica possibile e avere comunque la possibilità di registrare lo schermo – mentre l'utente verrà registrato mediante telecamere esterne.

I valutatori che dovranno osservare l'utente invece staranno nei suoi paraggi cercando di essere il più discreti possibili, ed intervenendo solo se strettamente necessario.

Se ciò dovesse avvenire, si prenderà prontamente nota di tale interferenza.

Metriche

Le variabili dipendenti sono i parametri che vengono presi in considerazione come misura dell'usabilità del sistema. Verranno utilizzate le seguenti metriche:

Per l'EFFICIENZA:

- Tempo di esecuzione del task

Per l'EFFICACIA

- Successo nello svolgimento del task

Per la SODDISFAZIONE

- Questionari somministrati agli utenti alla fine dello svolgimento di ciascun test, con domande aperte e affermazioni da valutare in un range da "Molto in disaccordo" a "Molto d'accordo".

Analisi dei Risultati

Nella fase di analisi dei risultati, si raccoglie l'intero materiale ottenuto con il test – questionari degli utenti, appunti dei valutatori, video – e li si esamina al fine di ottenere dei dati che possano essere realmente utili nella valutazione dell'applicazione testata.

Tali risultati verranno poi forniti allo sviluppatore del sistema, che potrà verificare quali sono stati i problemi riscontrati con maggiore frequenza, o che hanno causato il maggior numero di problemi, ed intervenire per risolverli stabilendo quindi una scala di priorità degli aspetti da affrontare per migliorare l'applicazione.

Oltre che mediante l'analisi dei risultati ottenuti dagli utenti, la valutazione verterà anche sulla verifica da parte dei tester del rispetto di alcune euristiche.

Dal materiale raccolto è possibile evidenziare se e quali problemi si sono verificati durante l'esecuzione dei task, con la relativa descrizione. Ad ogni problema riscontrato, viene assegnato un grado di priorità, che indica quanto è importante che quell'errore sia aggiustato il più presto possibile. Terremo conto soprattutto del numero di volte che si è riscontrato il problema, anche in relazione al livello di skill tecnologiche dell'utente che ha commesso l'errore; inoltre, riteniamo di particolare rilevanza anche capire se il verificarsi dell'errore ha in qualche modo inficiato sul risultato finale dell'esecuzione del task.

La priorità con cui devono essere risolti i problemi va da alta a bassa. I problemi con alta priorità vanno chiaramente risolti per primi.

Priorità alta: sono tutti quei problemi che devono essere risolti in maniera immediata, dato che non permettono di portare a termine il compito prefissatosi.

Priorità media: funzionalità secondarie del sistema, che potrebbero comunque aiutare di più l'utente.

Priorità bassa: riguardano i problemi minori di usabilità del sistema come ad esempio label poco chiare, interfaccia non responsive, ecc.

Come accennato nella sezione sulle metodologie da utilizzare per il testing, le euristiche da noi prese in esame per la valutazione sono le Euristiche di Nielsen.

Per ciascuna di esse verrà evidenziato come viene implementata dal sistema, o se al contrario essa non trova corrispondenza nell'applicazione.

1. Visibilità dello stato del sistema

- a. Tale euristica sembra essere pienamente rispettata: una volta selezionata ciascuna schermata, il nome della schermata sarà visualizzato sia in alto che sotto al bottone relativo al pannello selezionato.

2. Corrispondenza tra sistema e mondo reale

- a. L'utilizzo di icone che riprendono pienamente l'elemento che vogliono individuare (esempio: casa per la schermata home, pillola per la schermata relativa ai farmaci..) aiuterà l'utente a comprendere immediatamente cosa troverà in ciascuna schermata.
- b. Riteniamo molto positiva ed intuitiva la scelta dello sviluppatore di visualizzare il livello di medicinale ancora disponibile in maniera grafica, mostrando cioè il flacone che si svuota man mano che il farmaco in questione viene consumato.

3. Controllo e Libertà

- a. Da tutte le schermate dell'applicazione risulta possibile raggiungere tutte le altre schermate principali, utilizzando un sistema di pannelli e bottoni comune a molte applicazioni moderne. Tale elemento è valutato quindi positivamente.

4. Consistenza e standard

- a. Tutte le schermate mostrano la stessa organizzazione grafica, con alcuni elementi che rimarranno visibili per qualunque schermata (l'intestazione e la barra con i bottoni di selezione pannello)

5. Prevenzione dell'errore

- a. Una volta intrapresa ciascuna azione, è disponibile una sola scelta (oltre a quella di tornare indietro). Sarà dunque molto difficile che l'utente possa sbagliare nello svolgimento di un task, una volta che l'avrà iniziato.

6. Riconoscimento anziché ricordo

- a. Le schermate sono tutte molto chiare, schematiche e rispondenti a quello che ci si aspetta oggi da un'applicazione per smartphone. I periodi di apprendistato saranno molto brevi o nulli.

7. Flessibilità d'uso

- a. Il sistema non fornisce metodi d'uso alternativi o scorciatoie per utenti esperti. Sebbene ciò contravvenga questa specifica euristica, è pur vero che lo scopo principale del sistema, cioè il ricordare l'assunzione di farmaci a persone anziane o che devono assumerne grande quantità e l'utilizzo di una piattaforma mobile, lasciano poco spazio a modalità d'uso alternative, non essendo ad esempio possibile utilizzare comandi rapidi da tastiera.

8. Design ed estetica minimalista

- a. Il sistema è in linea con il *material design* definito da Google: Le schermate sono graficamente semplici e presentano pochi 'fronzoli' e nessun elemento non strettamente necessario all'uso.

9. Aiuto all'utente

- a. Nel prototipo fornitoci non sono previsti prompt di conferma per le azioni principali, né sembrano essere presenti messaggi di errore. Sebbene data la semplicità d'uso dell'interfaccia ciò non rappresenti necessariamente un problema, sarebbe bene che lo sviluppatore riesaminasse questo aspetto.

10. Documentazione

- a. Allo stato attuale dello sviluppo non sembra essere presente una documentazione. Anche in questo caso, l'estrema semplicità d'uso potrebbe non renderla necessaria. Tuttavia, lo sviluppatore dovrebbe comunque prendere in considerazione l'ipotesi di rendere disponibile una documentazione per il sistema.

Risultati dei test con gli utenti

Come illustrato precedentemente, gli utenti che hanno testato il sistema sono 10.

Si tratta, nel dettaglio, di 5 persone con skill tecnologiche di base e 5 con skill tecnologiche avanzate.

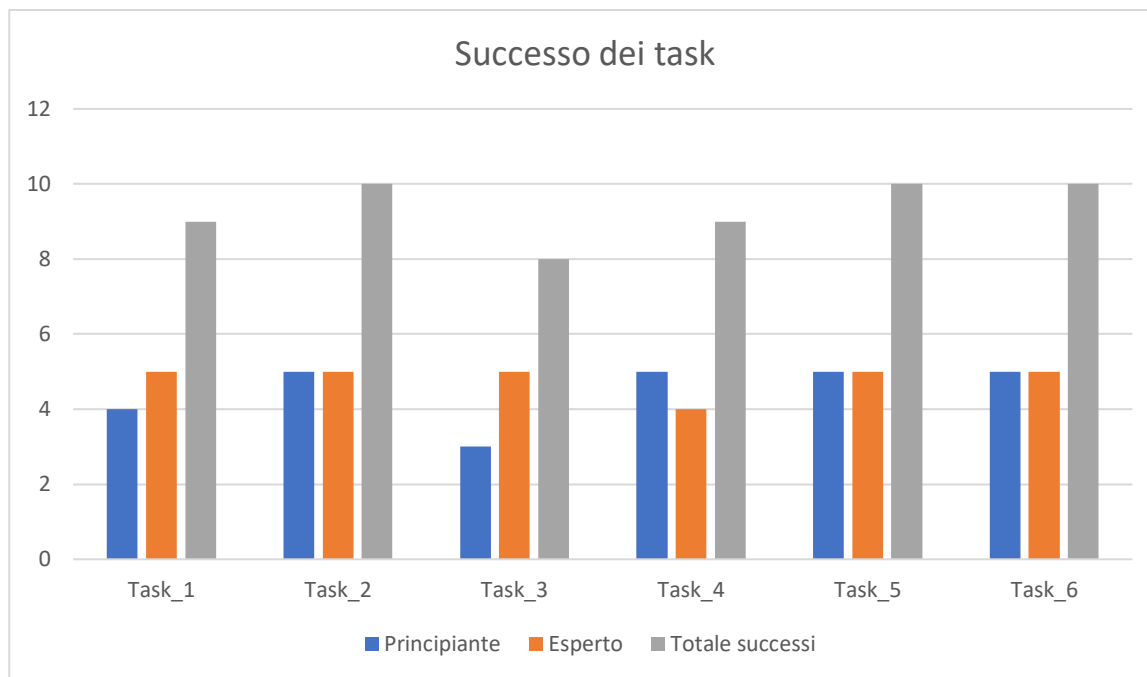
Successo dei test

Nei 10 test seguiti, vengono riportati quelli completati con successo (S), quelli completati parzialmente (P) e quelli non completati (F).

Tipologia utenti	Id utente	Task 1	Task 2	Task 3	Task 4	Task 5	Task 6
Esperto	Utente_1	S	S	S	S	S	S
Esperto	Utente_2	S	S	S	S	S	S
Esperto	Utente_3	S	S	S	P	S	S
Esperto	Utente_4	S	S	S	S	S	S
Esperto	Utente_5	S	S	S	S	S	S
	TotSuccessiEsperti	5	5	5	4	5	5
Principiante	Utente_6	S	S	S	S	S	S
Principiante	Utente_7	S	S	P	S	S	S
Principiante	Utente_8	S	S	F	S	S	S
Principiante	Utente_9	F	S	S	S	S	S
Principiante	Utente_10	S	S	S	S	S	S
	TotSuccessiPrincipianti	4	5	3	5	5	5
	TotSuccessi	9	10	8	9	10	10

Calcoliamo quindi il tasso di successo in relazione ai task falliti, ai task completati parzialmente e a quelli completati con successo. I task parziali vengono calcolati con la metà del punteggio di un successo.

$$\text{Tasso successo} = \frac{56 + (2 \cdot 0.5)}{60} = 95\%$$



Tempo impiegato

Nella seguente tabella viene illustrato il tempo impiegato per il completamento dei task (in secondi) ed il numero di errori che gli utenti hanno commesso per completare un task.

Tipologia utenti	Id utente	Task 1	Task 2	Task 3	Task 4	Task 5	Task 6	$\mu\mu_{Time}$
Esperto	Utente_1	16	19	23	20	20	21	
Esperto	Utente_2	12	21	20	21	17	19	
Esperto	Utente_3	19	25	19	27	20	19	
Esperto	Utente_4	22	16	21	19	19	22	
Esperto	Utente_5	20	21	20	18	23	18	
	μ_{Time}	17.8	20.4	20.6	21	19.8	19.8	19.9
Principiante	Utente_6	30	15	31	29	23	26	
Principiante	Utente_7	27	29	36	32	22	28	
Principiante	Utente_8	23	23	40	25	27	34	
Principiante	Utente_9	35	27	29	28	25	25	
Principiante	Utente_10	26	26	21	27	33	29	
	μ_{Time}	28.2	24	31.4	28.2	26	28.4	27.7
	$\mu_{TimeTotale}$	23	22.2	26	24.6	22.9	24.1	23.8 ($\mu\mu_{TimeTotale}$)

Il seguente grafico mostra il tempo impiegato al completamento dei task.



Numero errori

Nella seguente tabella viene illustrato il numero di errori compiuto da ciascun utente nello svolgimento dei task.

Tipologia utenti	Id utente	Task 1	Task 2	Task 3	Task 4	Task 5	Task 6
Esperto	Utente_1	0	0	0	1	0	0
Esperto	Utente_2	0	0	0	0	0	1
Esperto	Utente_3	1	0	0	1	0	0
Esperto	Utente_4	0	0	0	0	0	0
Esperto	Utente_5	0	0	0	0	0	1
	TotEsperti	1	0	0	2	0	2
Principiante	Utente_6	0	0	0	1	0	0
Principiante	Utente_7	0	0	2	0	0	1
Principiante	Utente_8	1	0	3	0	0	0
Principiante	Utente_9	5	0	0	0	0	1
Principiante	Utente_10	0	0	0	1	0	1
	TotPrincipianti	6	0	5	2	0	3
	Totale	7	0	5	4	0	5

Come è possibile evincere dalla tabella, il task su cui sono stati compiuti più errori è il task 1.

A seguire il task 3 e 6, ed infine il task 4.

Nello svolgimento del task 2 e 5 non si sono verificati errori.



Risultati operazioni tramite feedback dell'utente

Tramite i feedback riportati dagli utenti, sono state portate alla nostra attenzione le seguenti questioni

1. Sarebbe desiderabile che oltre al nome del farmaco venisse indicato anche il principio attivo. (Priorità media)
2. Sarebbe utile se, cliccando sul nome del farmaco, ne venisse visualizzato il bugiardo – dato che quest'ultimo è comunque sempre reperibile online. (Priorità media)
3. Sarebbe utile se nelle notifiche che ricordano di assumere un farmaco venisse visualizzato anche se tale farmaco sta finendo ed è necessario riacquistarlo. (Priorità media)
4. Sarebbe utile se nella schermata di riepilogo delle prossime visite fosse disponibile un bottone di chiamata rapida accanto al nome di ogni medico, studio medico o ospedale. (Priorità bassa)

Soddisfazione utenti

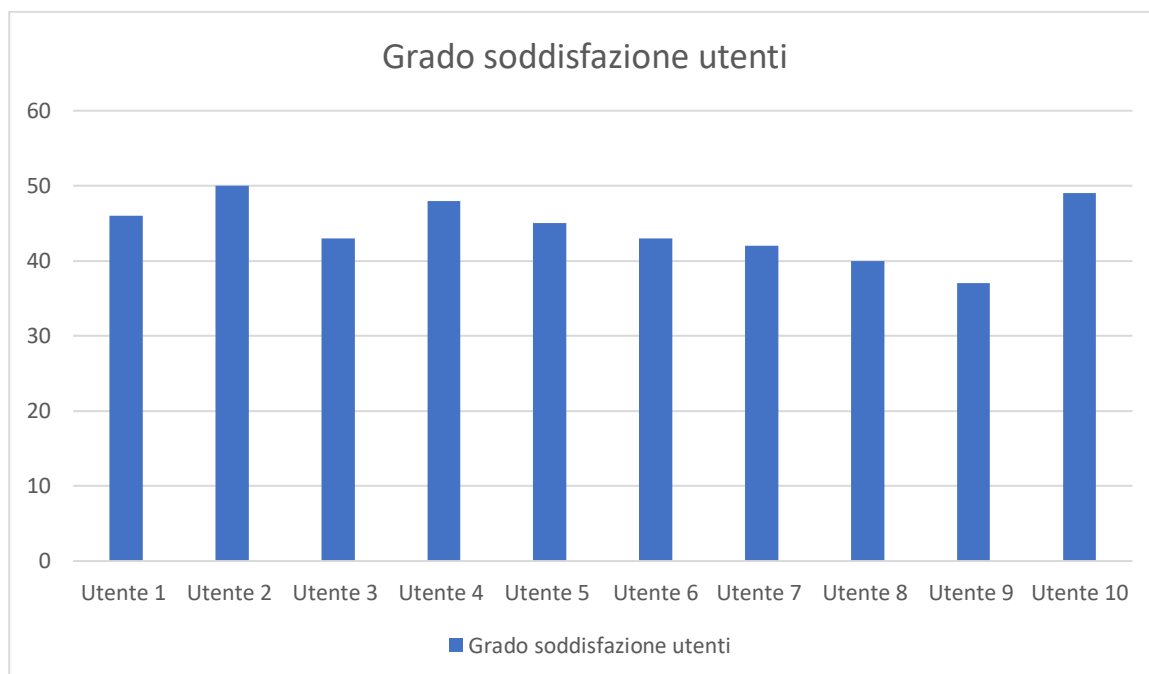
Tramite il questionario che segue la scala di Likert abbiamo rilevato il grado di soddisfazione dell'utente che ha usato l'applicativo. Il questionario riportato precedentemente consta di 10 domande (con un punteggio da 1 a 5, per un totale massimo di 50) chiuse e 3 domande aperte. Alle domande aperte non è attribuito punteggio, in quanto servono solo a capire se per l'utente il sistema potrebbe migliorare ulteriormente.

Riportiamo quindi un grafico contenente tutti punteggi dei test degli utenti.

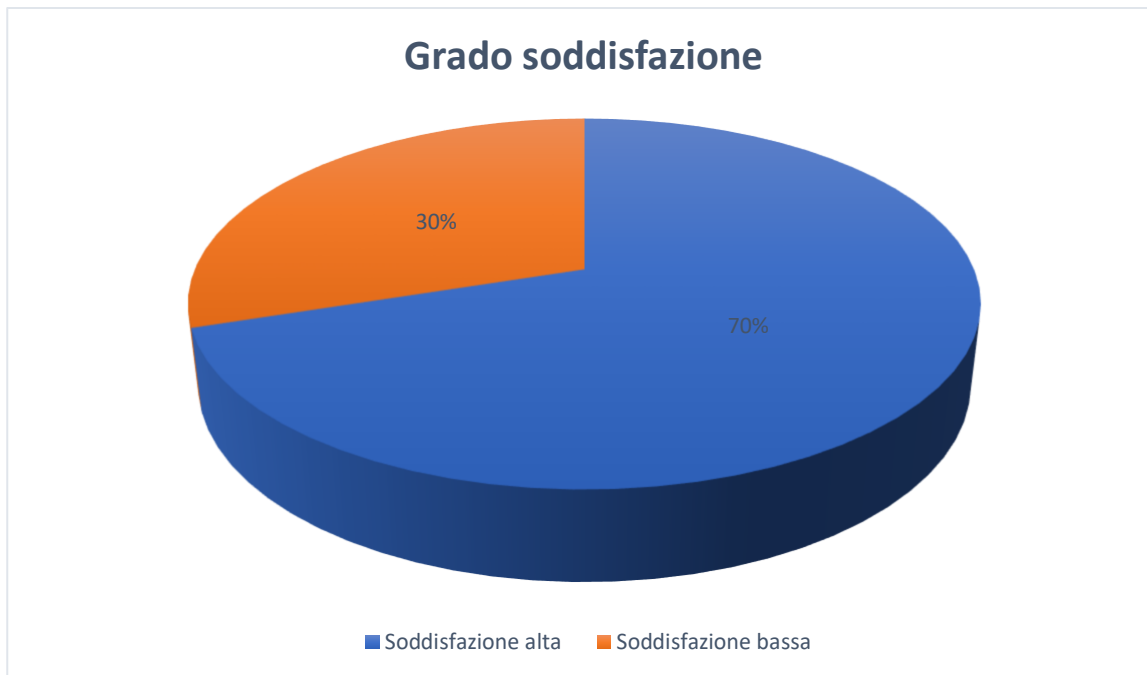
Utente	Tipo Utente	Punteggio
Esperto	Utente_1	46
Esperto	Utente_2	50
Esperto	Utente_3	43
Esperto	Utente_4	48
Esperto	Utente_5	45
Principiante	Utente_6	43
Principiante	Utente_7	42
Principiante	Utente_8	40
Principiante	Utente_9	37
Principiante	Utente_10	49

Il calcolo del punteggio è stato effettuato sommando le risposte date dagli utenti.

Nel seguente grafico a barre sono riportati tali punteggi.



Successivamente, questi valori sono riportati in percentuale - avendo effettuato la giusta proporzione - per ricavarne un grafico a torta. Gli utenti che hanno riportato un tasso di soddisfazione superiore all' 85% sono stati considerati molto soddisfatti, mentre quelli che hanno riportato un tasso inferiore sono stati considerati poco soddisfatti.



Invalidazione ipotesi nulla

Consideriamo l'ipotesi nulla H_{01} – *'non c'è differenza in termini di efficacia tra gli utenti che hanno skill tecnologiche avanzate e skill tecnologiche di base'*. È possibile osservare dai risultati ottenuti che il numero di errori e il tempo di completamento dei task risultino apprezzabilmente minori negli utenti che hanno skill tecnologiche avanzate; al contrario, risultano più alti negli utenti meno esperti con la tecnologia.

Tenendo conto dei risultati riportati negli appositi grafici e tabelle, l'ipotesi nulla risulta **invalidata**.

Essendo il sistema MedMinder pensato anche per utenti meno esperti è necessario che i progettisti lavorino al fine di renderlo, ove possibile, ancora più semplice, dal momento che il target sarà prevalentemente di persone anziane e tipicamente non abituate all'utilizzo di strumenti tecnologici.

Conclusioni

Il testing svolto ha previsto la suddivisione di 10 partecipanti in due categorie: utenti con skill tecnologiche avanzate e utenti con skill tecnologiche di base.

La scelta del gruppo in cui l'utente dovesse ricadere è stata effettuata mediante un questionario circa il background tecnologico.

Questo tipo di testing è stato quindi svolto per capire quanto effettivamente fosse semplice ed intuitiva l'interfaccia per tutte le categorie di utenti, che hanno potuto esprimere il proprio parere circa il grado di soddisfazione nell'uso del sistema mediante un *questionario di reazione* che è stato somministrato a fine test.

Per stabilire quanto l'applicazione fosse usabile abbiamo utilizzato come parametri il tempo di completamento dei task ed il numero di errori effettuati.

Analizzando successivamente i dati riportati, abbiamo notato come il tempo di completamento ed il numero di errori commessi dipendessero dal background tecnologico: più gli utenti hanno skill avanzate, più è alta la probabilità che un task sia portato a termine con successo in un lasso di tempo minore e con meno errori.

Lo studio di usabilità effettuato fin ora è replicabile, in quanto ripetendo i test con persone diverse appartenenti alle due categorie di skill tecnologiche, nelle stesse condizioni, si otterrebbero risultati molto simili tra loro.

Ciò è dovuto principalmente al fatto che sono state usate delle metriche oggettive, quali tempo di completamento e numero di errori effettuati.

Con lo studio di usabilità effettuato abbiamo quindi coperto tutte le questioni di usabilità da individuare, in quanto è risultato che:

1. il sistema è efficace, poiché permette di portare a termine tutti i task che ci eravamo prefissati di testare;
2. il sistema è efficiente, poiché il tempo di completamento dei task ed il numero di errori commessi non risultano eccessivamente elevati, neanche per utenti poco esperti con la tecnologia;
3. il grado di soddisfazione dell'applicazione è molto elevato, in quanto il 70% degli utenti è molto soddisfatto dall'interazione con il sistema.

Dai risultati ottenuti, l'interfaccia non si è rivelata troppo complicata per gli utenti, e le uniche problematiche individuate hanno priorità di risoluzione medio/bassa.

Detto ciò, visto che il target dell'applicazione saranno principalmente utenti anziani o con poca esperienza tecnologica, non si può non tenere conto delle differenze – sebbene non eccessive – tra efficacia ed efficienza di utenti con skill tecnologiche e quelle di utenti privi di tali skill.

Ci dichiariamo quindi soddisfatti dei risultati dell'usabilità dell'applicativo MedMinder, suggerendo allo stesso tempo allo sviluppatore di esaminare nuovamente l'interfaccia grafica proposta, cercando ove possibile di renderla ancora più intuitiva.

Riferimenti

1. <http://www.far.unito.it/usabilita/Cap5.htm>
2. Slide del corso di Interazione Uomo Macchina
3. Mary Beth Rosson and John M. Carroll - Usability Engineering. Scenario-Based Development of Human-Computer Interaction (2001, Morgan Kaufmann)