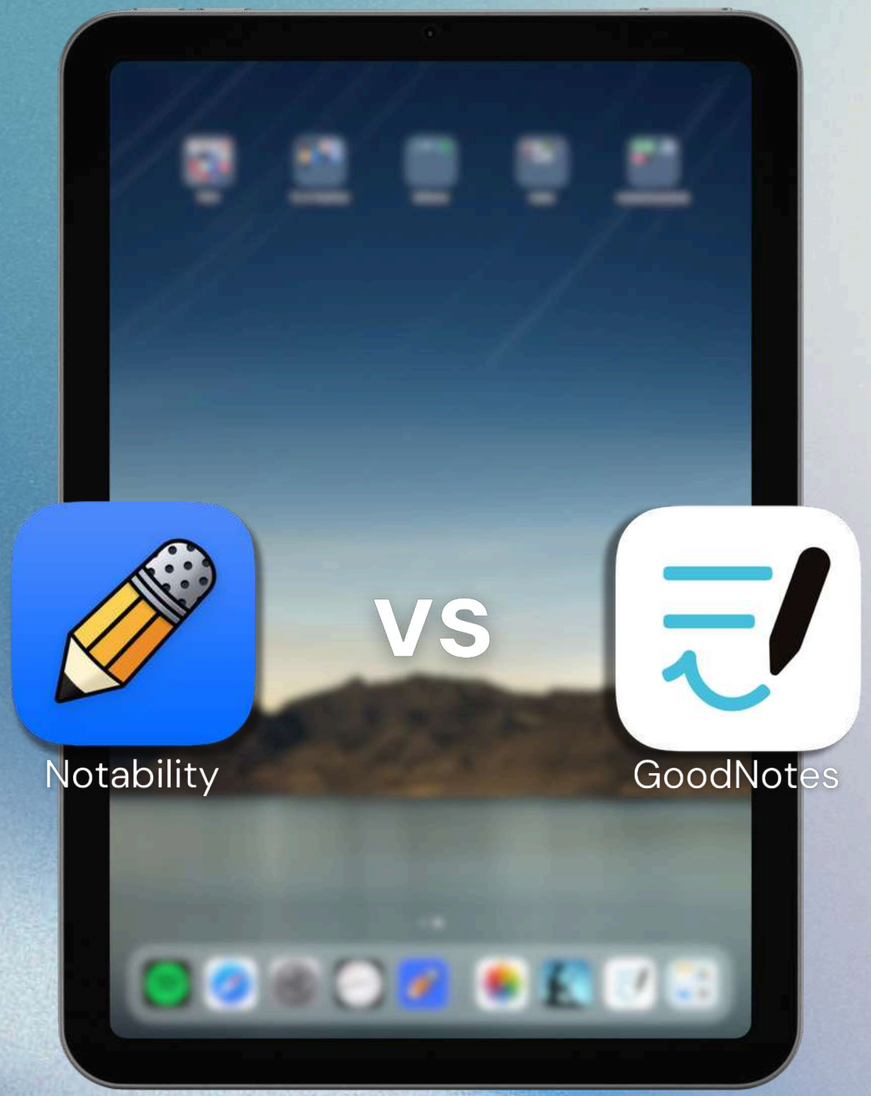
Analisi di Usabilità

Confronto Trasversale



A CURA DI:

- De Marchi Sara, 879188
- Ferrara Luigi, 879222
- Leskovec Ethan, 886040
- Romaniello Andrea, 890316

INDICE

1 Introduzione

2 Valutazione Euristica

3 Test Utente

4 Questionario

5 Conclusioni

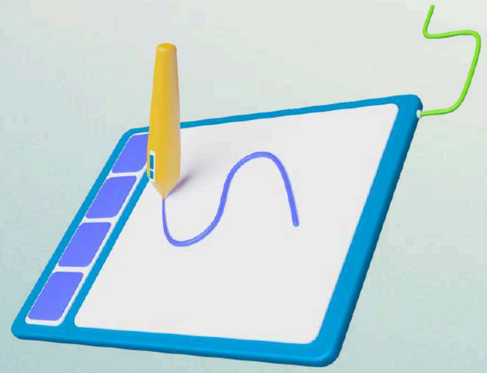
6 Appendice

INTRODUZIONE

Introduzione al problema

GoodNotes

Notability



Ambiente di valutazione

Introduzione al Problema

L’utilizzo di **tecnologie** come **supporto** in ambienti istituzionali e didattici è una prassi ormai comune.
In particolare, se ci focalizziamo su contesti accademici come quelli universitari, è essenziale considerare strumenti che permettano agli studenti di seguire e prendere appunti in maniera dinamica durante e dopo le lezioni.

A questo proposito, l’utilizzo di **tablet** è fondamentale. Questi ultimi infatti consentono di annotare informazioni, leggere libri e testi digitali, salvare slides e presentazioni in modo semplice e sfruttare un ambiente di apprendimento migliore per un’assimilazione più veloce delle informazioni.

In questo studio mettiamo a confronto due applicazioni che permettono di svolgere i compiti sopra citati, consideriamo l’usabilità di quest’ultime e valutiamo se effettivamente **portano a una massimizzazione della produttività** umana.

Le applicazioni in questione sono: **Notability** e **GoodNotes**.

GoodNotes è nata per **migliorare l'esperienza degli appunti tradizionali su carta**, offrendo funzionalità avanzate come la possibilità di prendere appunti digitali scritti a mano, la ricerca di testo scritto a mano e l'organizzazione dei documenti in una libreria digitale personalizzata. Attualmente, GoodNotes sta lavorando per **integrare l'intelligenza artificiale** nella scrittura a mano digitale, portando l'innovazione a un nuovo livello.

GoodNotes

Fondata in Australia nel **2011**, GoodNotes è disponibile per sistemi iOS, Android, Windows, Mac e in versione Web.

Nel **2022** è stata premiata da Apple come **migliore app dell'anno per iPad**.

Notability

Creata nel 2010, Notability è l'applicazione per prendere appunti **più venduta sull'App Store di Apple**.

Notability è disponibile per iPad, Mac e iPhone

Notability è stata progettata per **facilitare la presa di appunti digitali** durante lezioni o riunioni di lavoro. Consente di svolgere attività di base come scrivere a mano su tablet, ma propone anche **funzionalità avanzate** come la registrazione audio sincronizzata automaticamente con gli appunti.

Ambiente di Valutazione

Valutazione Mobile

Per la realizzazione del progetto è stato svolto il confronto trasversale tra le due applicazioni nelle loro versioni mobile. È inoltre stato utilizzato lo stesso sistema operativo, **iPadOS**, e **Apple Pencil** di prima e seconda generazione.

Versione

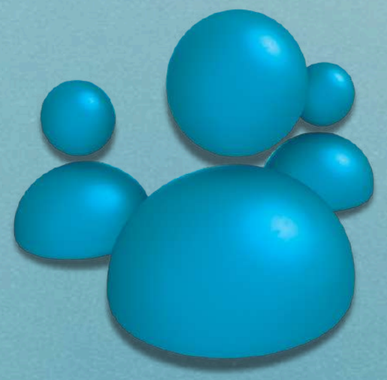
Le due applicazioni sono state utilizzate nelle loro versioni base, ovvero **Notability Classic** e **GoodNotes 5**.

Account

Gli utenti hanno utilizzato le due applicazioni con **account preesistenti**, in modo da potersi concentrare esclusivamente sui task preparati per loro.

VALUTAZIONE EURISTICA

Overview
ed Euristiche



Valutatori

Problemi
riscontrati

Conclusioni

Overview

In questa prima parte proponiamo un’**analisi qualitativa dell’usabilità** attraverso una valutazione euristica svolta da **utenti esperti**.

Lo scopo di questa valutazione è quello di identificare il **maggior numero possibile di criticità** e **problemi di usabilità** delle due applicazioni sotto esame.

Dopo la valutazione tali problemi verranno analizzati rispetto ad un insieme di **principi di buona progettazione** e ordinati in base alla priorità e alla severità.

Sono riportate inoltre alcune **frasi rilevanti** registrate durante le **valutazioni olistiche** e i **dark pattern*** riscontrati in ciascuna delle due applicazioni.

Euristiche

Il set di princìpi di riferimento scelto per l'analisi sono le **10 euristiche di Nielsen**, riportate qui a fianco.

*Un'**euristica** è un insieme di concetti, riferimenti e strategie che si sono rilevate particolarmente adatte a risolvere un determinato problema che, nel nostro caso, è la **progettazione di sistemi interattivi usabili.***

ID	Euristica
E1	Visibilità dello stato del sistema
E2	Corrispondenza tra sistema e mondo reale
E3	Controllo e libertà dell'utente
E4	Assicurare consistenza
E5	Riconoscimento piuttosto che ricordo
E6	Assicurare flessibilità ed efficienza
E7	Visualizzare tutte e sole le informazioni necessarie
E8	Prevenire gli errori
E9	Aiuto nel rilevare e correggere gli errori
E10	Aiuto e documentazione

PERCEZIONE COGNIZIONE ERRORE

Valutatori

Per la valutazione qualitativa sono stati coinvolti 7 valutatori, che hanno svolto in ordine:

- Una **valutazione di tipo olistico****, durante la quale è stato richiesto di esprimere ad alta voce cosa stessero pensando mentre esploravano le due applicazioni, cosa piacesse o meno e cosa trovassero più o meno facile da utilizzare (questo processo è anche conosciuto come *Think-Aloud Protocol Analysis*).
- Dopo aver stilato una lista dei problemi identificati, è stato chiesto ai valutatori di assegnare un **grado di severità** a tali problemi, in modo da poterli poi ordinare per priorità.

Valutazione Olistica: l'utente esplora il sistema liberamente, in maniera non guidata

In questa prima fase il campione è composto da:

4 ESPERTI DI USABILITÀ*

*(corrispondenti ai quattro membri del gruppo che hanno realizzato questo progetto)

3 ESPERTI DI DOMINIO

ID UTENTE	GENERE	FASCIA D'ETÀ	OCCUPAZIONE	FAMILIARITÀ
EU1	M	<25	Studente	Alta
EU2	M	<25	Studente	Alta
EU3	F	<25	Studente	Alta
EU4	M	<25	Studente	Media
ED1	M	<25	Studente	Alta
ED2	M	<25	Studente	Media
ED3	F	<25	Studente	Media

EU: Esperto di Usabilità **ED: Esperto di Dominio**

Matrice di Expertise

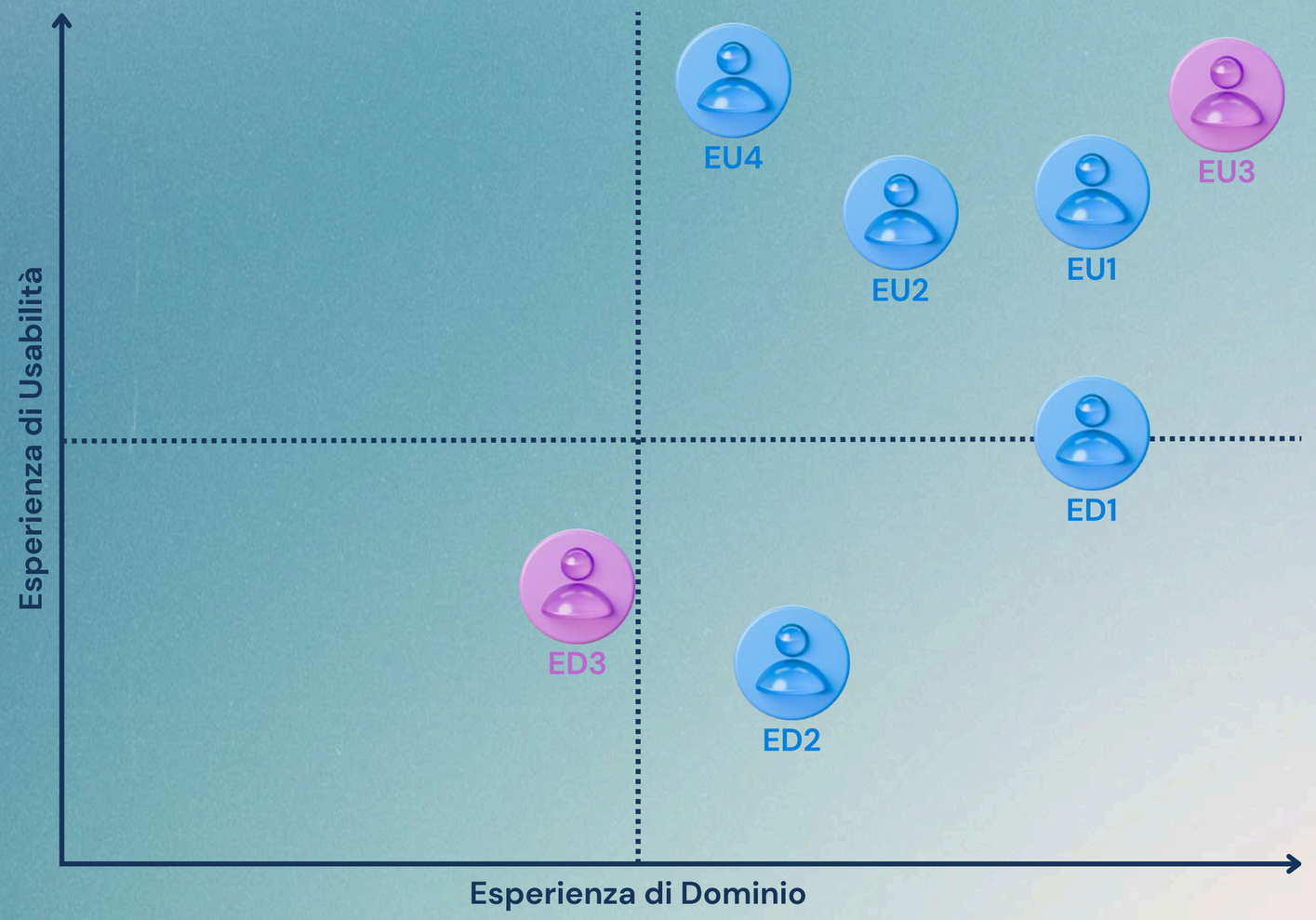
La matrice di expertise permette di vedere

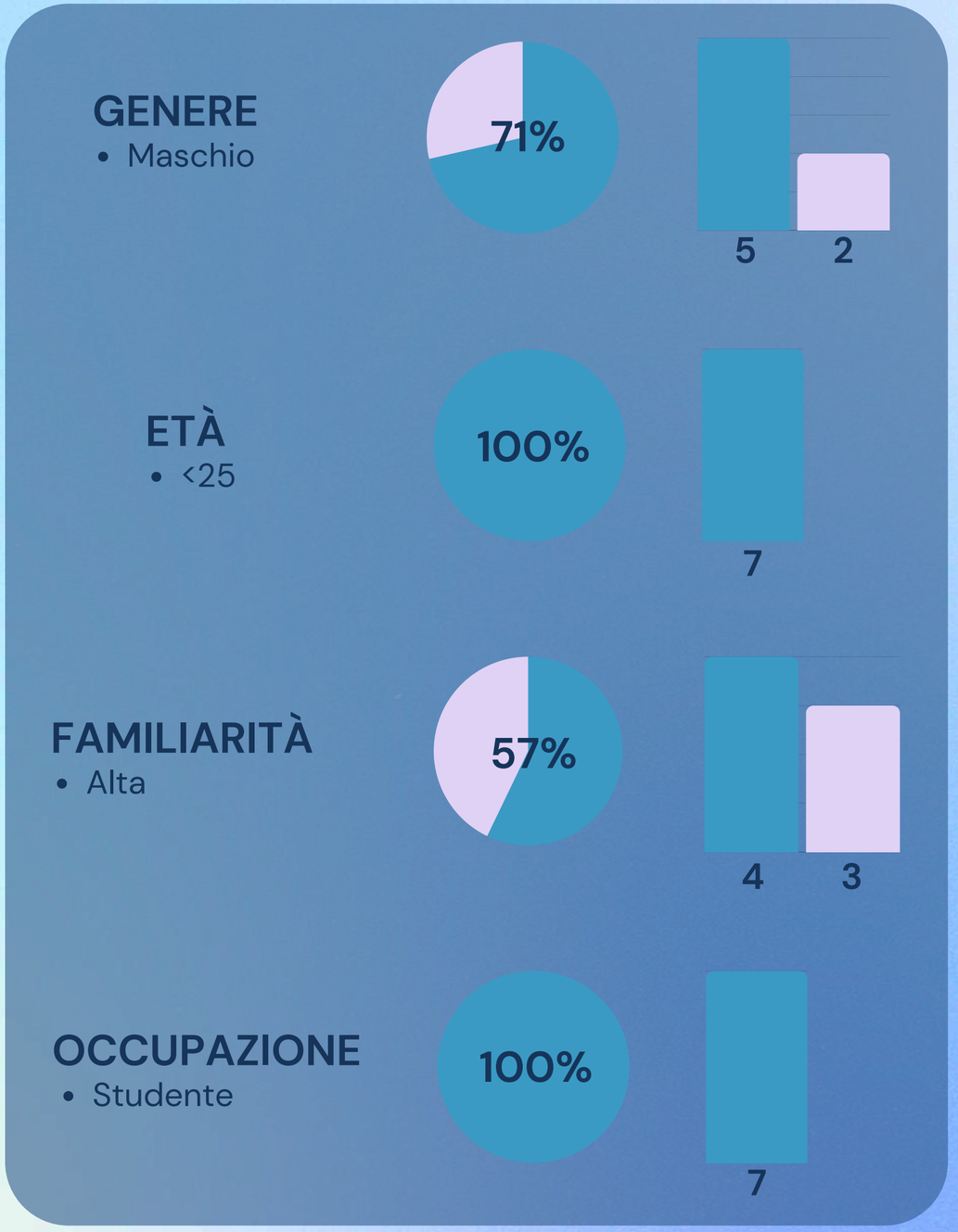
- l'**esperienza di usabilità**
- l'**esperienza di dominio**

di ciascun utente coinvolto nella valutazione euristica.

In particolare nel nostro caso si può osservare:

- una conoscenza medio-alta del dominio e un'alta conoscenza dell'usabilità da parte dei 4 esperti di usabilità;
- una minore conoscenza dell'usabilità ma una conoscenza medio-alta del dominio per quanto riguarda i 3 esperti di dominio.





Composizione del campione

I valutatori coinvolti sono:

- prevalentemente di sesso **maschile**
- tutti di età **inferiore ai 25 anni**
- perlopiù con **familiarità alta**
- tutti **studenti**.

Problemi riscontrati

Nelle prossime slides verranno presentati i problemi riscontrati durante la valutazione euristica.
In particolare saranno presentati:

- Criteri di Prioritizzazione
- Tabella dei Problemi Riscontrati
- Problemi più Rilevanti
- Distribuzione delle Euristiche
- Matrice Problemi-Valutatori
- Dark Patterns
- Frasi Rilevanti

Criteri di prioritizzazione

I valutatori hanno assegnato a ciascun problema un valore tra 0 e 4 in base a 3 parametri:

- **Frequenza del problema**
- **Impatto del problema**
- **Persistenza del problema**

Grazie alle valutazioni dei valutatori coinvolti è stato possibile dividere il problema in **3 fasce di priorità**.
È stata calcolata la frequenza con cui ogni problema è stato valutato tra i più gravi, utilizzando il Principio di Pareto*, selezionando quindi il 20%.
Sono stati ricavati **2 gruppi di problemi**:

SIGNIFICATO DEI VALORI

0 *= Non è un problema di usabilità.*
1 *= Da risolvere solo in caso di tempo extra.*
2 *= Problema secondario, con bassa priorità.*
3 *= Problema rilevante, con alta priorità.*
4 *= Problema di assoluta priorità, imperativo risolverlo!*

GRUPPO 1
Problemi indicati
più frequentemente
come tra i più gravi

GRUPPO 2
Problemi indicati
meno frequentemente
come tra i più gravi

*il Principio di Pareto afferma che circa il 20% delle cause provoca l'80% degli effetti.

Criteri di prioritizzazione

Successivamente, è stato svolto un **One Proportion Test** per capire se la **differenza** tra il numero di problemi presenti nel GRUPPO 1 e nel GRUPPO 2 fosse **statisticamente significativa** o dovuta al caso.
In base ai risultati, i problemi sono stati ulteriormente divisi in **tre fasce di priorità*** in base alla loro **significatività**:

FASCIA A
- Priorità Alta

FASCIA B
- Priorità Media

FASCIA C
- Priorità Bassa

I risultati ottenuti sono riportati nelle slide seguenti in forma tabellare ordinate per:

Nelle tabelle sono inoltre riportati ID del problema, descrizione breve e completa, le euristiche che sono state violate, da quanti e quali valutatori è stato identificato e valore medio, mediano e deviazione standard della severità.

CRITERI DI ORDINAMENTO:
1. **Classe di Priorità** (da A a C)
2. **Valore di Severità media** (decrescente)
3. **Valore di Popolarità** (decrescente)

*Fasce A, B e C sono state calcolate utilizzando il python script chiamato **gravity2priority**

Problemi riscontrati - GoodNotes

ID PROBLEMI	DESCRIZIONE BREVE	DESCRIZIONE COMPLETA	ID ESURISTICHE VIOLATE	POPOLARITÀ	ID VALUTATORI	PRIORITÀ	SEVERITÀ MEDIA (SD)	SEVERITÀ MEDIANA
PB5	Funzione "Lente" non facilmente reperibile	L'icona che rappresenta la funzione lente non viene individuata facilmente.	E3, E6	4	EU1, EU2, EU4, ED2	A	2.7142 (0.6998)	3
PB18	Salvataggio selezione	Quando si utilizza lo strumento "Lazo" è possibile salvare l'elemento selezionato nella propria raccolta di sticker, ma la voce che permette di svolgere tale funzione ("Aggiungi elemento") non è significativa di ciò.	E1, E6	2	ED1, ED2	B	2.5713 (1.4983)	3
PB11	Troppe funzioni	La barra degli strumenti presenta un numero eccessivo di funzioni che possono talvolta disorientare. Le funzioni posso essere raggruppate in maniera migliore	E6, E8	1	EU4	B	2.4286 (0.4949)	2
PB14	Font e grandezza nella funzione "Tastiera"	La funzione "Tastiera" permette di scrivere un testo sul documento. Non è possibile cambiare il font e la grandezza del suddetto testo.	E3, E6, E7	1	EU4	B	2.4286 (0.4949)	2

Problemi riscontrati - GoodNotes

ID PROBLEMI	DESCRIZIONE BREVE	DESCRIZIONE COMPLETA	ID ESURISTICHE VIOLATE	POPOLARITÀ	ID VALUTATORI	PRIORITÀ	SEVERITÀ MEDIA (SD)	SEVERITÀ MEDIANA
PB4	Bottone "Seleziona" non chiaro	L'icona della funzione "seleziona" (affianco ad" impostazioni") non è rappresentativa.	E2	1	EU3	B	2.1428 (0.9897)	3
PB10	Problema tastiera nome cartella	Alla creazione di una nuova cartella, se si sta utilizzando la penna per tablet la tastiera non appare automaticamente ma va aperta cliccando un bottone a parte	E1, E2, E5, E6	2	EU2, ED2	B	2 (1.1952)	3
PB15	Funzione "Zoom" non chiara	La funzione "Zoom" ingrandisce una zona del documento permettendo di scrivere in maniera più precisa. L'icona che identifica tale funzione non è rappresentativa.	E2, E5, E6	1	EU2	C	2.1429 (0.6389)	2
PB3	Ambiguità creazione documento	Ambiguità nel processo di creazione di un documento. La funzione crea taccuino ha un nome ambiguo	E4, E5	2	EU2, EU3	C	2.1428 (1.1248)	2

Problemi riscontrati - GoodNotes

ID PROBLEMI	DESCRIZIONE BREVE	DESCRIZIONE COMPLETA	ID ESURISTICHE VIOLATE	POPOLARITÀ	ID VALUTATORI	PRIORITÀ	SEVERITÀ MEDIA (SD)	SEVERITÀ MEDIANA
PB1	Posizione funzione “eliminati di recente”	La funzione cestino si trova in una posizione non ottimale	E2, E6, E9	1	EU4	C	2 (0.5345)	2
PB9	Funzione “Lazo” non chiara	Non è chiaro come utilizzare lo strumento “Lazo”. In particolare nelle situazioni di copia e incolla.	E1. E6	2	EU1, EU2	C	2 (1.069)	2
PB19	Pressione penna inutile	E’ possibile modificare la pressione della penna ma non apporta nessuna modifica alla stessa	E3	1	ED1	C	1.8571 (1.3553)	2
PB2	Funzioni ridondanti	Ambigua differenza tra la funzione ‘Goodnotes 6’ e la funzione ‘store’. Potrebbero essere accorpate in un’unica funzione	E4, E8	3	EU1, EU3, EU4	C	1.5714 (0.9035)	1

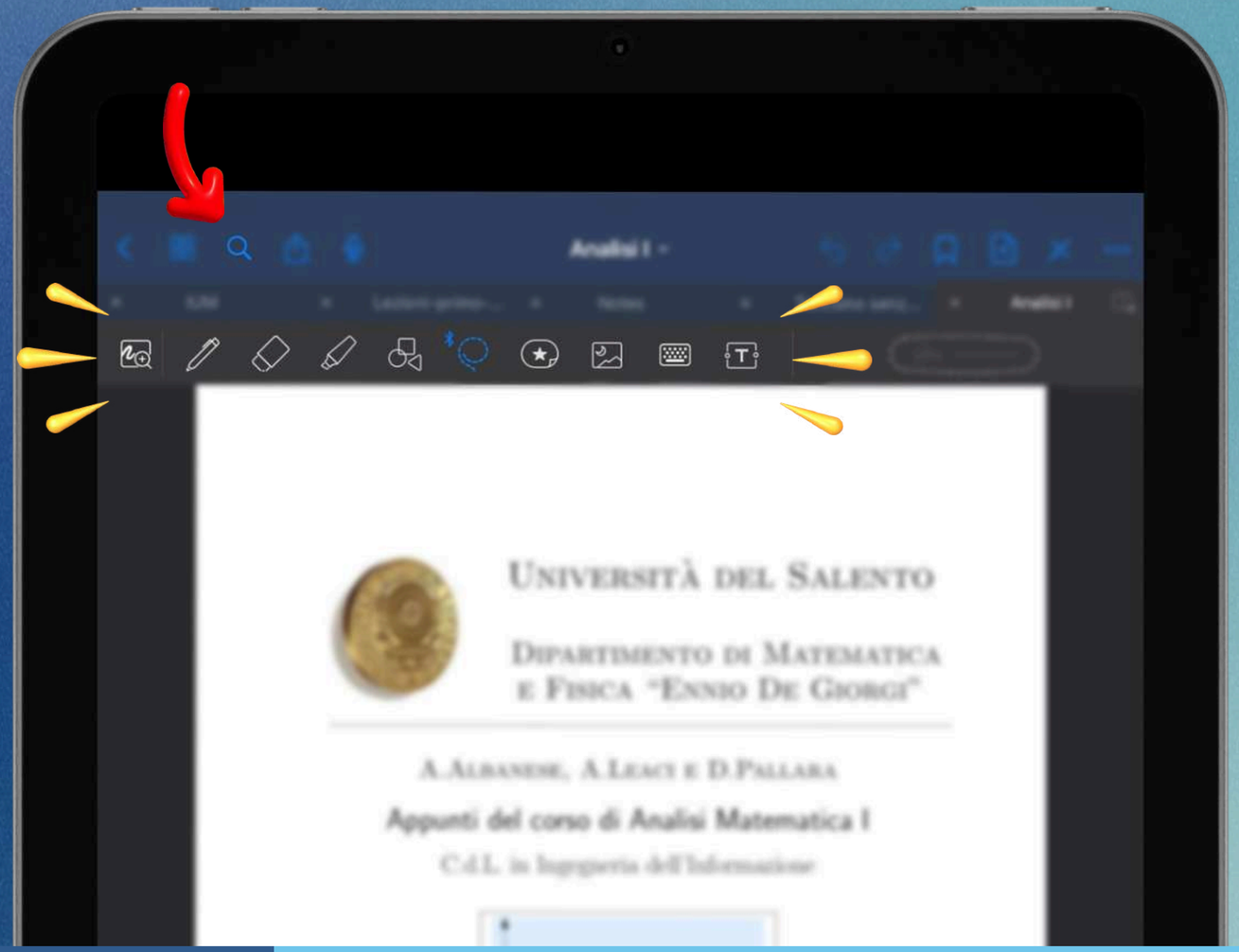
Problemi riscontrati - GoodNotes

ID PROBLEMI	DESCRIZIONE BREVE	DESCRIZIONE COMPLETA	ID ESURISTICHE VIOLATE	POPOLARITÀ	ID VALUTATORI	PRIORITÀ	SEVERITÀ MEDIA (SD)	SEVERITÀ MEDIANA
PB8	Doppia voce di “Backup”	Nelle impostazioni esistono due voci “Backup dei dati” e “Backup automatico” che potrebbero essere raggruppate in un’unica voce.	E4, E6, E8	2	EU3, EU4	C	1.5714 (0.7284)	1
PB17	Modifica tratto penna	Non è ben chiaro come fare per modificare il tratto della penna.	E5, E6, E7	1	ED1	C	1.5714 (0.7284)	1
PB16	Ridondanza bottoni scrittura	Le funzioni “tastiera” e “casella di testo” potrebbero essere accorpate in un unico bottone invece che due separati.	E7, E8	2	EU3, EU4	C	1.4286 (0.9035)	1
PB6	Icona “v” non chiara	L’icona “v” a fianco del nome del documento non è significativa. Non è rappresentativa della funzione che svolge.	E2, E5	1	EU3	C	1.2857 (0.6999)	1
PB7	Icona “Segnalibro” non chiara	L’icona “Segnalibro” in alto a destra in ogni documento non è rappresentativa della funzione che svolge.	E1	1	EU4	C	1 (1.069)	1

Problemi riscontrati - GoodNotes

ID PROBLEMI	DESCRIZIONE BREVE	DESCRIZIONE COMPLETA	ID ESURISTICHE VIOLATE	POPOLARITÀ	ID VALUTATORI	PRIORITÀ	SEVERITÀ MEDIA (SD)	SEVERITÀ MEDIANA
PB12	Bottone “sticker” non significativo	Il bottone della funzione “sticker” non è rappresentativo	E2, E5, E6	2	EU3, EU4	C	1 (0.5345)	1
PB32	Ridimensionamento elementi	Non è possibile ridimensionare gli elementi “pizzicando” con le dita; bisogna usare i pulsanti appositi.	E5	1	ED3	C	1 (1.069)	1
PB13	Bottone “tastiera” non significativo	Non è chiaro cosa faccia il bottone con l’icona della tastiera.	E6	1	EU2	C	0.8571 (0.6389)	1
PB33	Indici poco intuitivi	Gli indici delle pagine sono poco intuitivi; non si capisce come utilizzarli.	E1, E7	2	ED2, ED3	C	0.8571 (0.6389)	1

Problemi più rilevanti - GoodNotes

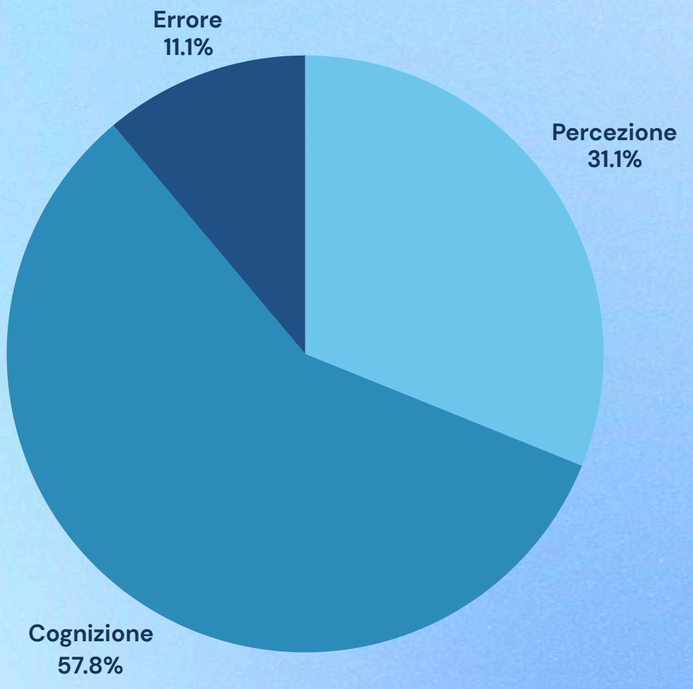


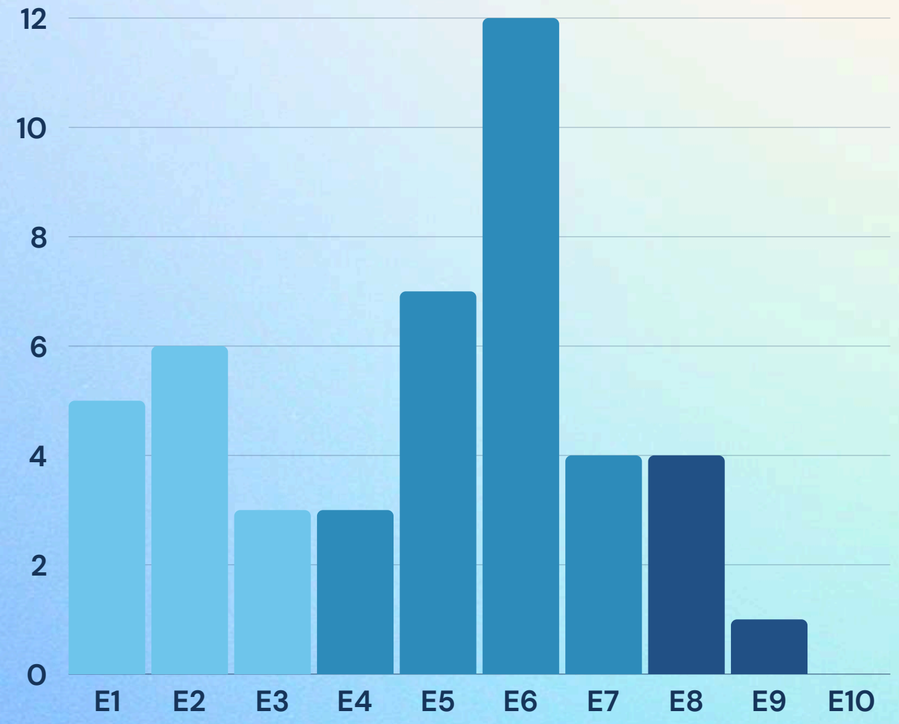
PROBLEMA PB5 –
Funzione "Lente" non facilmente reperibile

L'icona che rappresenta la funzione lente non viene individuata facilmente.

Gli utenti hanno difficoltà nella ricerca dello strumento "Lente" all'interno dell'interfaccia. In particolare sembra che li disorienti il fatto che non si trovi nella barra dove stanno tutti gli altri strumenti principali.

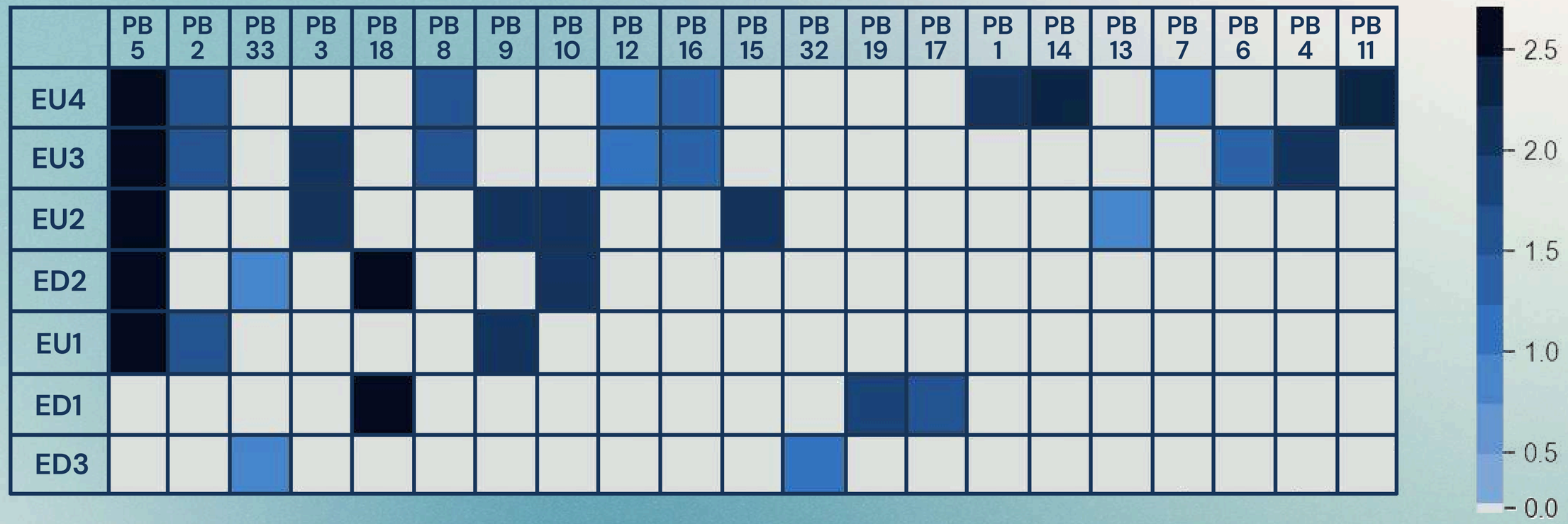
Distribuzione euristiche-problemi





Dall'analisi dei problemi riscontrati risulta che le **euristiche di cognizione** vengano violate con **maggiore frequenza** rispetto alle altre. In particolare l'**euristica 6** risulta essere **la più violata**.

Matrice problemi-valutatori



La matrice mette in relazione i **problemi identificati** nella fase di valutazione euristica, con i **valutatori che li hanno riscontrati**. La matrice è stata ordinata in modo da avere i problemi trovati con più frequenza a sinistra e i valutatori migliori in alto.

Ad ogni problema è stato inoltre assegnato un **colore** che ne indica il **grado di severità** (secondo la scala a fianco della matrice).

Considerazioni

Visti i dati raccolti, è possibile approssimare il numero stimato degli errori in totale, con la seguente formula:

$$\frac{x \times y}{z}$$

con:
x = numero di errori trovati dal primo esaminatore
y = numero di errori trovati dal secondo esaminatore
z = numero di errori trovati da entrambi gli esaminatori

Nell'analisi sono stati scelti:

- **EU4** come **PRIMO ESAMINATORE**, essendo quello con il maggior numero di problemi trovati (= 9)
- **EU3** come **SECONDO ESAMINATORE**, essendo quello con il minor numero di problemi trovati (= 8)

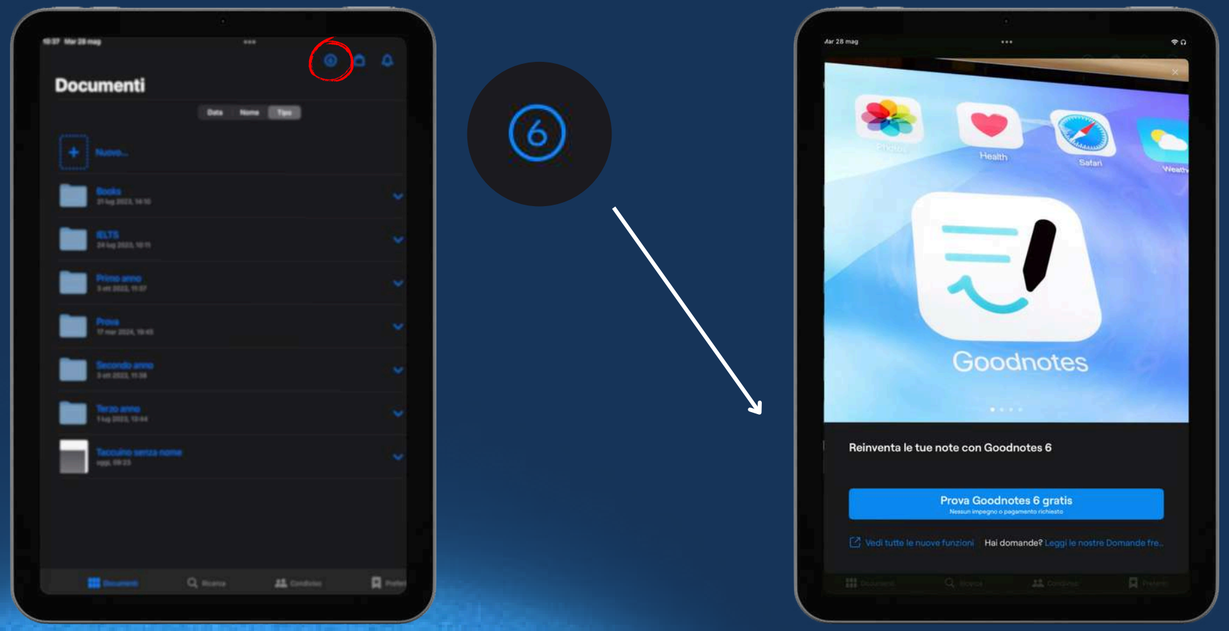
Utilizzando la formula precedentemente riportata, avendo 5 problemi rilevati da entrambi i valutatori, otteniamo un totale stimato di 14 errori.

Avendo individuato 21 problemi durante la valutazione euristica, quindi in **numero superiore a quelli stimati**, è possibile quindi assumere che:
SIA STATA INDIVIDUATA ALMENO LA TOTALITÀ DEI PROBLEMI DI USABILITÀ PER L'APPLICAZIONE DI GOODNOTES.

Dark Patterns

I Dark Patterns sono delle **scelte di design** nei sistemi, implementate con l'obbiettivo di **disorientare** l'utente; a volte per **ingannarlo** a sottoscrivere abbonamenti che non voleva, altre invece per **manipolarlo psicologicamente**.

Dark Patterns - GoodNotes

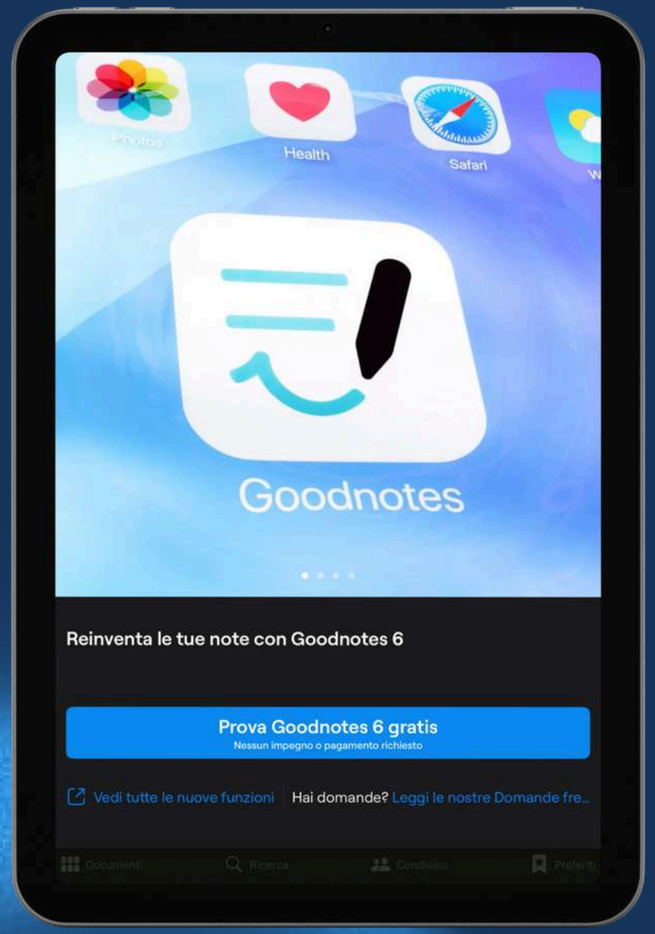


DISGUISED ADS

*"Le **pubblicità** sono **nascoste nella pagina**, come se fossero parte del "normale" contenuto"*

Su GoodNotes, l'Icona "6" è insieme ad altre funzionalità, come le notifiche, invece è una pubblicità per effettuare l'upgrade alla versione GoodNotes 6.

Dark Patterns - GoodNotes

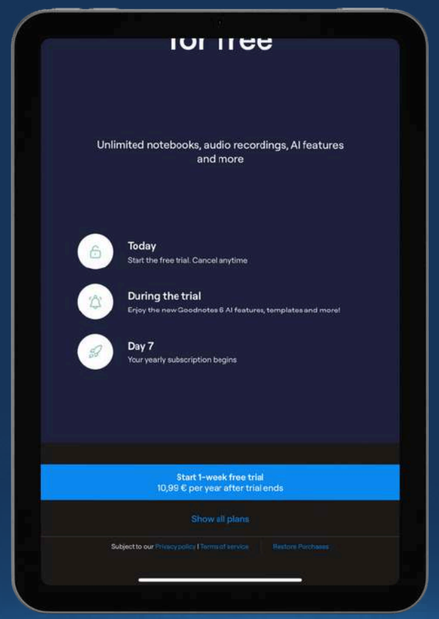


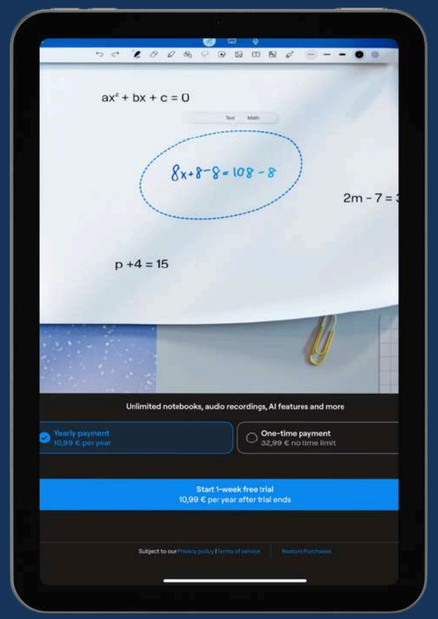
FORCED CONTINUITY

*"Tecnica manipolativa in cui un servizio richiede informazioni di pagamento per un **periodo di prova gratuito** per poi **attivare automaticamente l'abbonamento a pagamento senza avvisare** chiaramente l'utente"*

GoodNotes forza il rinnovo automatico sia per l'abbonamento annuale che per il passaggio dalla prova gratuita alla versione completa, senza nessuna notifica precedente.

Dark Patterns - GoodNotes



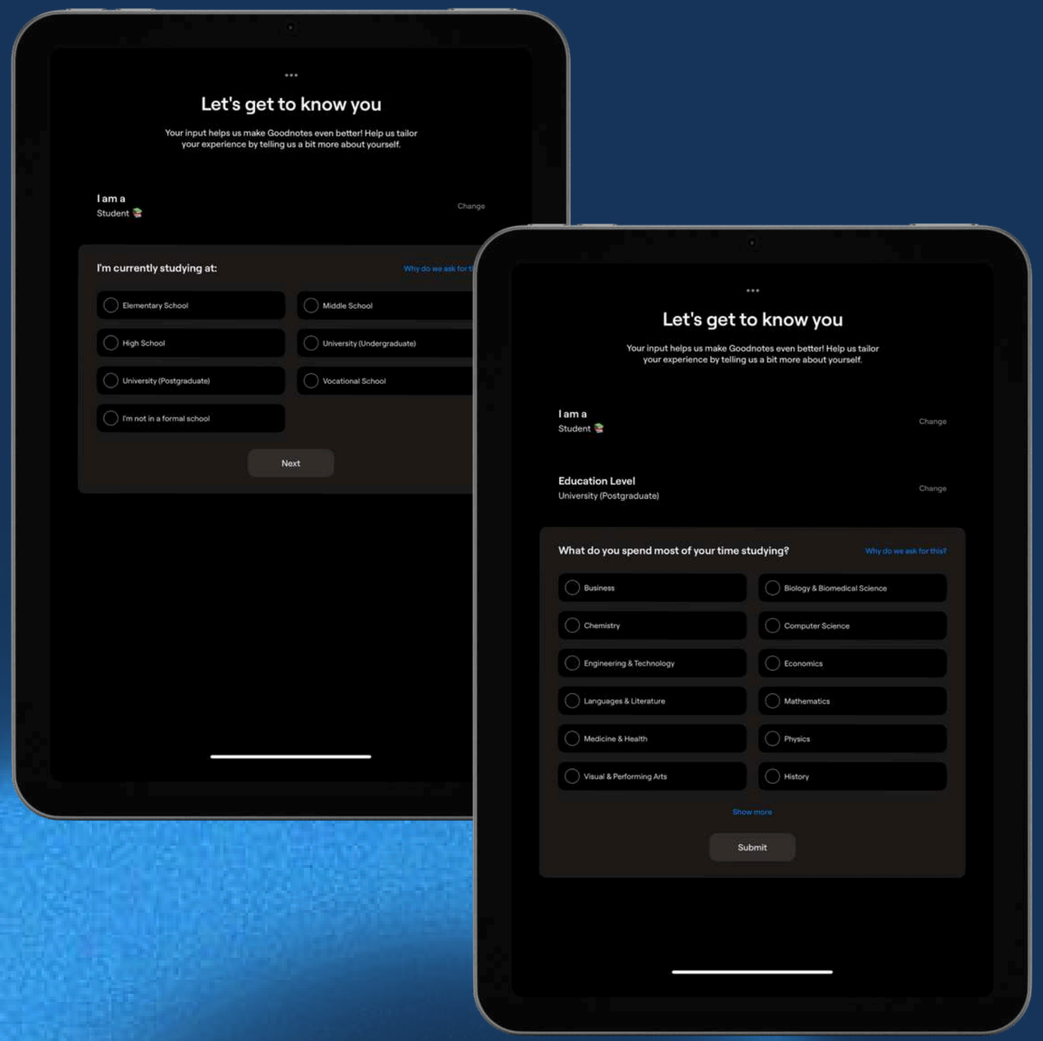


PRICE COMPARISON PREVENTION

*"Il sistema rende **difficile** per l'utente **comparare il prezzo** con quello di un altro prodotto, in modo da non fargli fare una decisione informata"*

GoodNotes nasconde l'opzione di pagamento one-time-only dietro un opzione nascosta, ostacolando l'utente nella scelta del piano a lui più congeniale.

Dark Patterns - GoodNotes



PRIVACY ZUCKERING

*"L'utente è "ingannato" nel **fornire più informazioni di sé stesso** di quante avesse davvero intenzione di condividere"*

GoodNotes rende impossibile uscire da questo processo una volta iniziato; vengono chieste informazioni come l'occupazione e molti ulteriori dettagli per completare il profilo.

Frasi Rilevanti - GoodNotes

Registrate durante la Valutazione Olistica degli Esperti di Dominio

"Non esiste l'impostazione... **Questo è un bel problema!**"

(in riferimento all'impostazione per cambiare il tratto della penna)

"Non si capisce un ca***"

(in riferimento alla funzione "Tastiera")

"Adegua lo stile... non adegua lo stile"

(frase registrata durante l'esplorazione delle opzioni disponibili una volta selezionato un elemento con lo strumento "Lazo")

"**Questa è una funzione interessante**, non c'è in tutte le applicazioni"

(in riferimento alla presenza di tab per navigare più facilmente tra i documenti aperti)

Problemi riscontrati - Notability

ID PROBLEMI	DESCRIZIONE BREVE	DESCRIZIONE COMPLETA	ID ESURISTICHE VIOLATE	POPOLARITÀ	ID VALUTATORI	PRIORITÀ	SEVERITÀ MEDIA (SD)	SEVERITÀ MEDIANA
PB26	Ambiguità icona “Pagine”	L’icona “Pagine” usa l’icona standard usata per le funzioni di “Incolla” o “Duplica”.	E2, E3, E4, E6	2	EU3, EU4	A	3.1428 (1.0669)	3
PB20	Bottone “Nuovo” non intuitivo	Il bottone “Nuovo” non propone la creazione di una nuova cartella, ma apre direttamente un nuovo documento vuoto, in una posizione non definita.	E3, E4, E5, E6, E8	2	EU4, ED3	A	2.8571 (0.6389)	3
PB18	Opzione “Salva” dello strumento “Lazo”	Quando si utilizza lo strumento “Lazo” è possibile salvare l’elemento selezionato come sticker. Tale funzione è indicata solo con la voce “Salva” che non da’ però alcuna indicazione su dove la selezione verrà salvata.	E1, E6	3	EU1, EU3, EU4	A	2.5714 (0.7284)	3
PB24	Ambigua differenza tra Argomento e Divisore	“Argomento” indica una cartella, “Divisore” indica un raccoglitore di “Argomenti”, di fatto un’altro tipo di cartella.	E2, E3, E5, E6, E8	2	EU3, ED3	B	2.5714 (0.9035)	33

Problemi riscontrati - Notability

ID PROBLEMI	DESCRIZIONE BREVE	DESCRIZIONE COMPLETA	ID ESURISTICHE VIOLATE	POPOLARITÀ	ID VALUTATORI	PRIORITÀ	SEVERITÀ MEDIA (SD)	SEVERITÀ MEDIANA
PB21	Dicitura “Argomento” non chiara	“Argomento” si riferisce a una generica “Cartella”	E2, E5	3	EU3, EU4, ED3	B	2.4286 (0.7284)	2
PB23	Funzione “Galleria” non chiara	La “Galleria” non è una galleria, è uno store di template per documenti da scaricare	E2, E4	1	EU1	C	2 (0.9258)	2
PB30	Ricerca documento	La ricerca di una parola restituisce una serie di documenti ma non propone un’anteprima della posizione di tale parola all’interno dei suddetti.	E1, E6, E7, E8	1	ED1	C	2 (1.069)	2
PB29	Scelta rapida impostazioni documento	Una volta creato un nuovo documento vengono proposte sullo schermo delle impostazioni di modifica dello stesso (carta a quadretti, righe, possibilità di importare o scansionare un documento). Se si seleziona una di tali opzioni le altre poi scompaiono e per modificare il documento è necessario aprire le impostazioni.	E3, E7, E9	1	ED1	C	1.8571 (0.9897)	2

Problemi riscontrati - Notability

ID PROBLEMI	DESCRIZIONE BREVE	DESCRIZIONE COMPLETA	ID ESURISTICHE VIOLATE	POPOLARITÀ	ID VALUTATORI	PRIORITÀ	SEVERITÀ MEDIA (SD)	SEVERITÀ MEDIANA
PB1	Posizione funzione “Cancellati di recente”	La posizione dei file cancellati di recente non è accessibile direttamente ma è inserita nella voce “Impostazioni”.	E2, E6, E9	1	EU4	C	1.5714 (0.9035)	2
PB15	Funzione “Zoom” non chiara	La funzione “Zoom” ingrandisce una zona del documento permettendo di scrivere in maniera più precisa. L’icona che identifica tale funzione non è rappresentativa.	E2, E5, E6	2	EU1, EU2	C	1.5714 (0.9035)	2
PB27	Funzione “Nastrino” non chiara	La funzione serve per creare delle zone interattive che si possono coprire e scoprire con un tocco.	E2, E5, E7	2	EU4, ED3	C	1.5714 (1.178)	2
PB28	Bottone aggiunta immagini non rappresentativo	Il bottone per l’aggiunta di immagini permette in realtà di aggiungere anche altri elementi visivi (come sticker e gif) ma dall’icona ciò non è chiaro.	E2, E5, E7	2	EU3, EU4	C	1.5714 (0.7284)	1

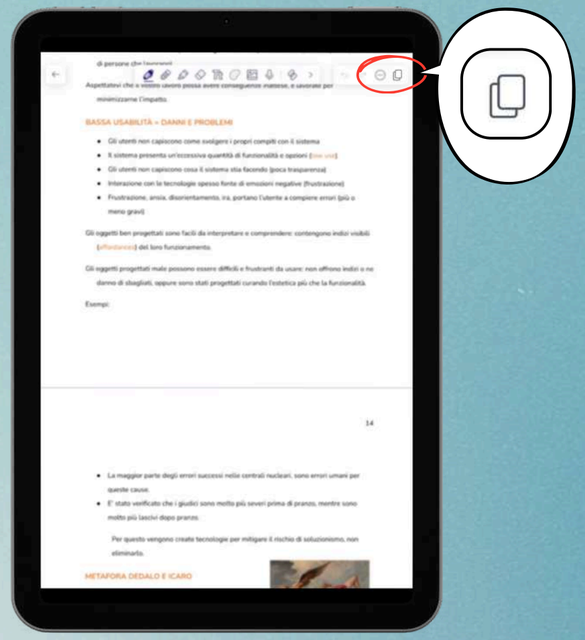
Problemi riscontrati - Notability

ID PROBLEMI	DESCRIZIONE BREVE	DESCRIZIONE COMPLETA	ID ESURISTICHE VIOLATE	POPOLARITÀ	ID VALUTATORI	PRIORITÀ	SEVERITÀ MEDIA (SD)	SEVERITÀ MEDIANA
PB31	Barra strumenti mobile	La barra degli strumenti non può essere fissata e può quindi essere spostata involontariamente mentre si sta modificando il documento.	E3, E8	1	ED1	C	1.4286 (1.2936)	2
PB9	Funzione “Lazo” non chiara	Non è chiaro come utilizzare lo strumento “Lazo”. In particolare nelle situazioni di copia e incolla.	E1, E6	2	EU1, EU2	C	1.2857 (1.2778)	1
PB22	Icona con i tre palllini non intuitiva	L’icona con i tre palllini di fianco a “Nuova” non da opzioni ulteriori per la suddetta funzione, ma offre funzioni di formattazione per i documenti	E3, E5	1	EU2	C	1.2857 (0.8806)	2
PB25	Posizione opzioni di condivisione	Le opzioni per la condivisione del documento non sono accessibili direttamente dalla barra degli strumenti ma sono inserite in una voce a parte.	E2, E6	1	EU4	C	1.2857 (1.0302)	1
PB32	Ridimensionamento elementi	Non è possibile ridimensionare gli elementi “pizzicando” con le dita; bisogna usare i pulsanti appositi.	E5	1	ED3	C	0.5714 (1.1338)	0

Problemi più rilevanti - Notability

Problema PB26 - Ambiguità icona “Pagine”

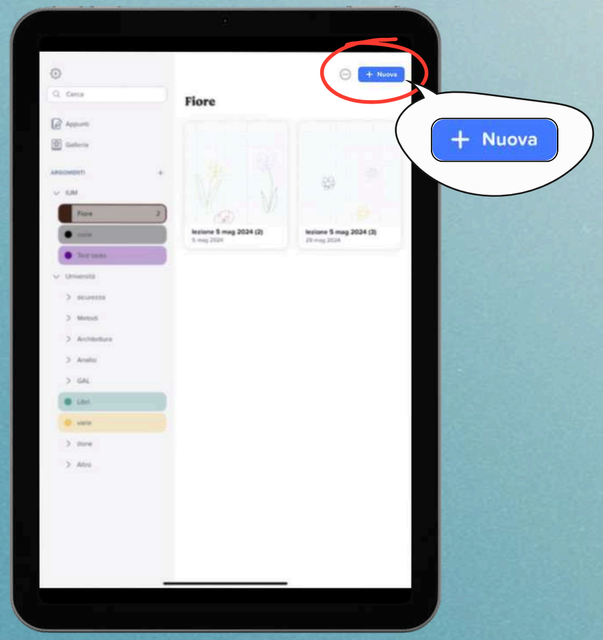
L’icona “Pagine” usa l’icona standard usata per le funzioni di “Incolla” o “Duplica”.



Gli utenti danno per scontato che l’icona che indica lo strumento per l’interazione con le pagine del documento indichi in realtà uno strumento per la funzione di copia e incolla.

Problema PB20 - Bottone “Nuovo” non intuitivo

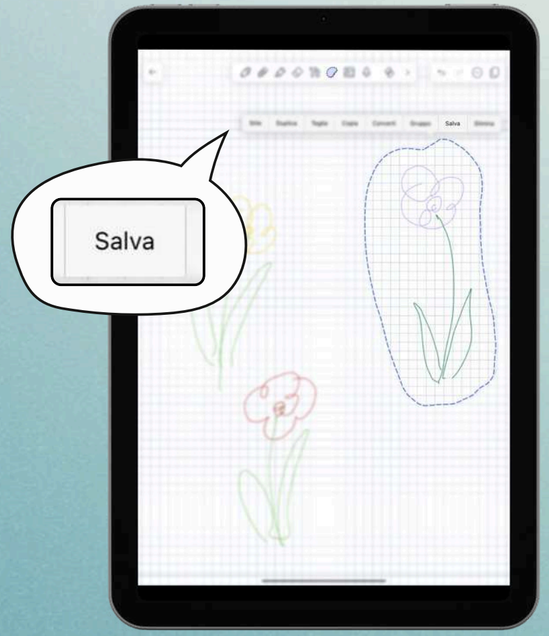
Il bottone “Nuovo” non propone la creazione di una nuova cartella, ma apre direttamente un nuovo documento vuoto, in una posizione non definita.



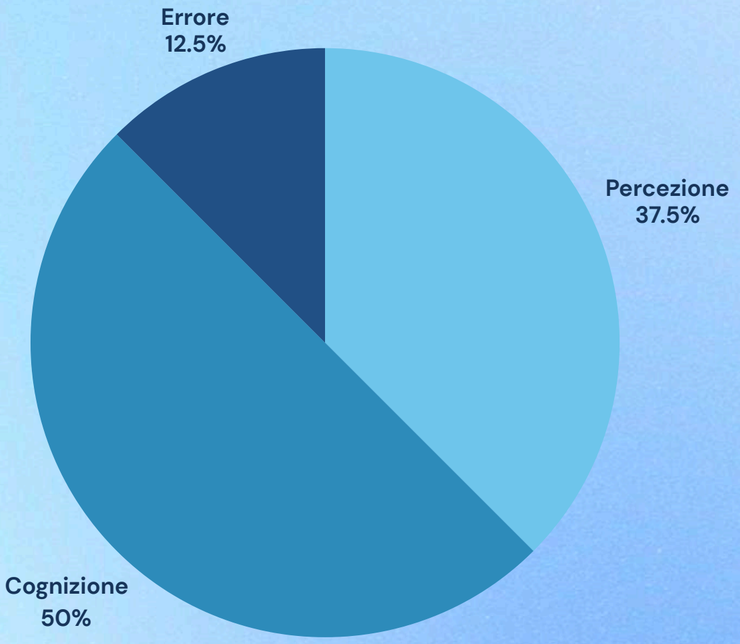
Gli utenti vengono “ingannati” dal tasto “Nuovo” che permette di creare solo un nuovo documento (ciò non viene indicato e gli utenti lo usano anche quando vogliono creare altro, ad esempio una cartella).

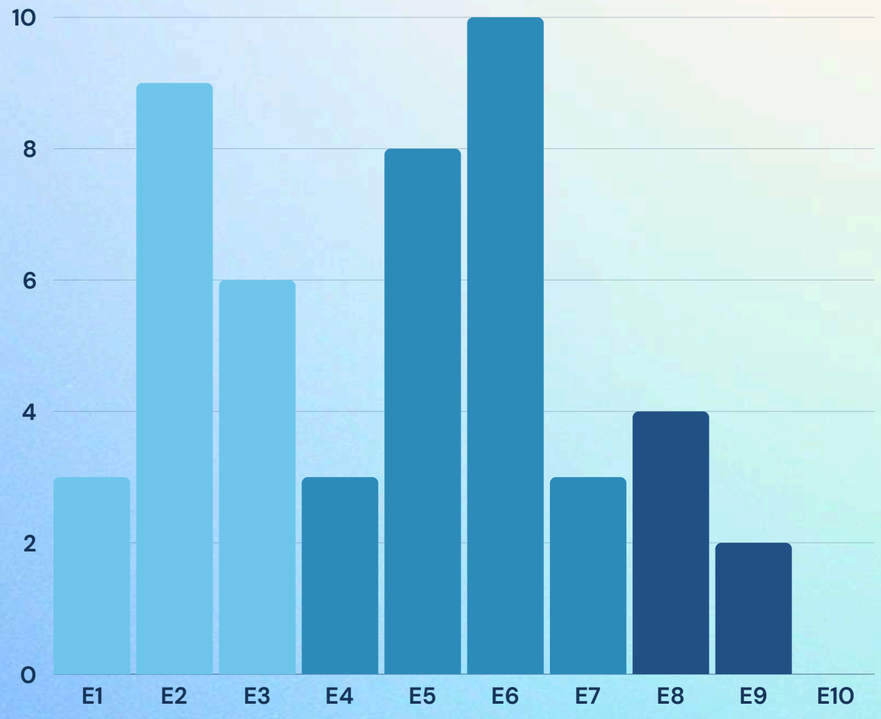
Problema PB18 - Opzione “Salva” dello strumento “Lazo”

Quando si utilizza lo strumento “Lazo” è possibile salvare l’elemento selezionato come sticker. Tale funzione è indicata solo con la voce “Salva” che non da’ però alcuna indicazione su dove la selezione verrà salvata.



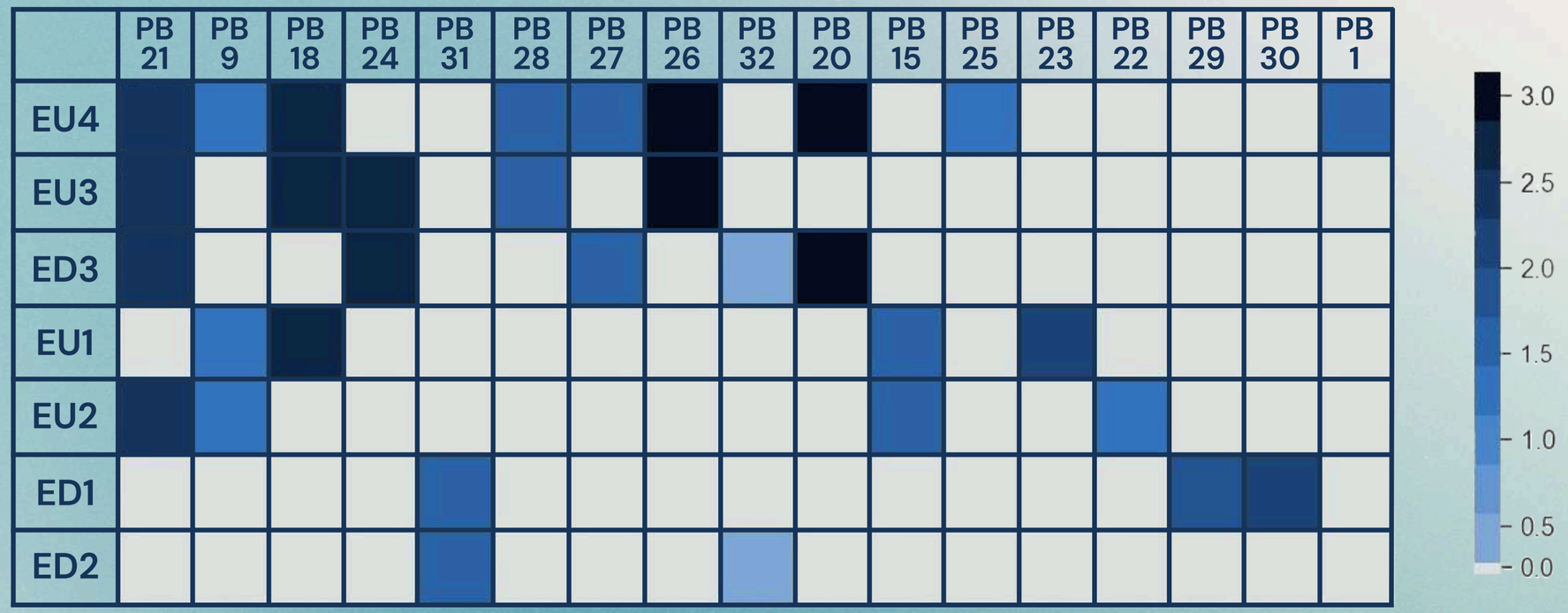
Distribuzione euristiche-problemi





Dall'analisi dei problemi riscontrati risulta che le **euristiche di cognizione** vengano violate con **maggiore frequenza** rispetto alle altre. In particolare l'**euristica 6** risulta essere **la più violata**.

Matrice problemi-valutatori



La matrice mette in relazione i **problemi identificati** nella fase di valutazione euristica, con i **valutatori che li hanno riscontrati**. La matrice è stata ordinata in modo da avere i problemi trovati con più frequenza a sinistra e i valutatori migliori in alto.

Ad ogni problema è stato inoltre assegnato un **colore** che ne indica il **grado di severità** (secondo la scala a fianco della matrice).

Considerazioni

Visti i dati raccolti, è possibile approssimare il numero stimato degli errori in totale, con la seguente formula:

$$\frac{x \times y}{z}$$

con:
x = numero di errori trovati dal primo esaminatore
y = numero di errori trovati dal secondo esaminatore
z = numero di errori trovati da entrambi gli esaminatori

Nell'analisi sono stati scelti:

- **EU4** come **PRIMO ESAMINATORE**, essendo quello con il maggior numero di problemi trovati (= 9)
- **EU2** come **SECONDO ESAMINATORE**, essendo quello con il minor numero di problemi trovati (= 5)

Utilizzando la formula precedentemente riportata, avendo 4 problemi rilevati da entrambi i valutatori, otteniamo un totale stimato di 11 errori.

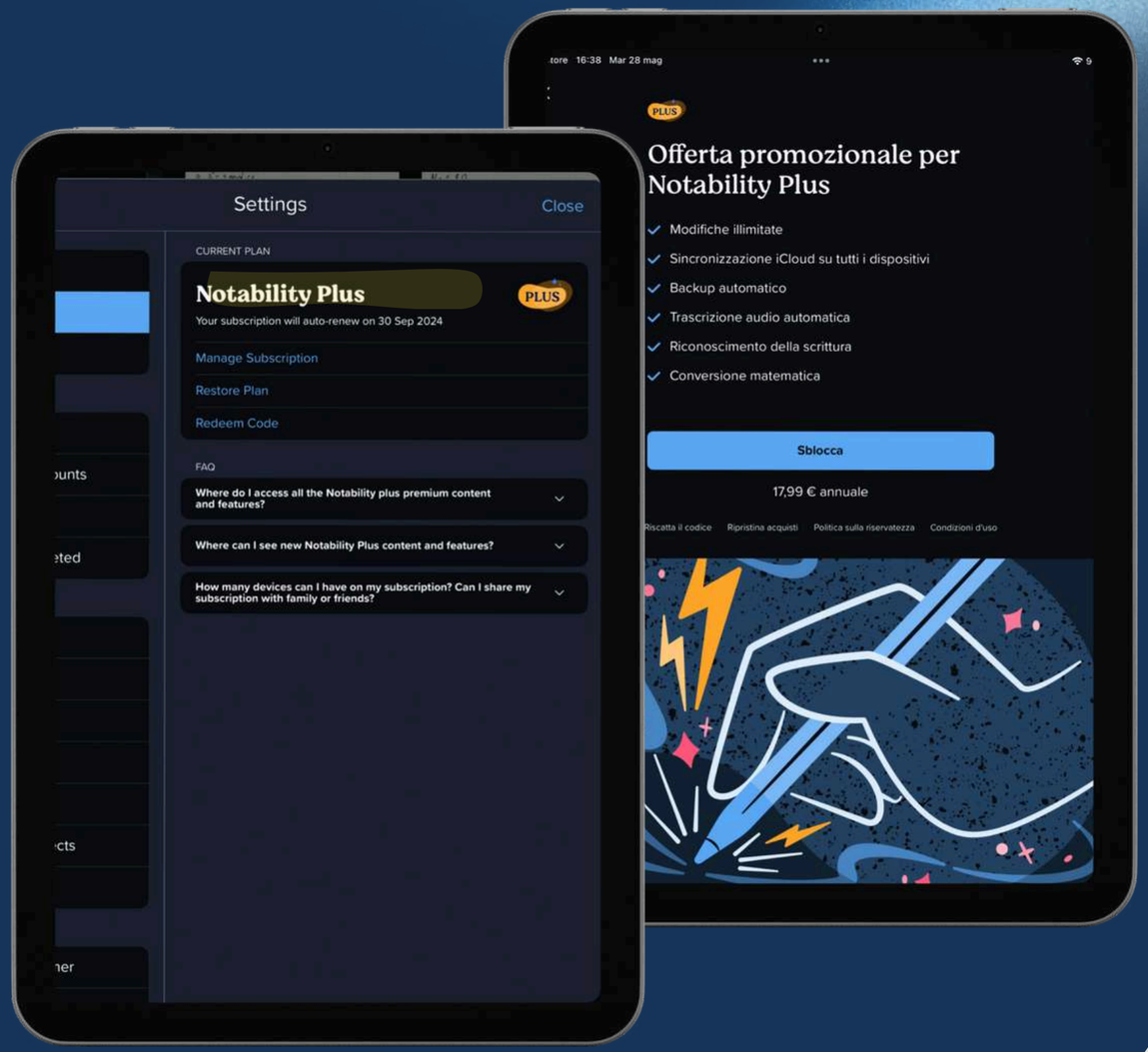
Avendo individuato 17 problemi durante la valutazione euristica, quindi in **numero superiore a quelli stimati**, è possibile quindi assumere che:
SIA STATA INDIVIDUATA ALMENO LA TOTALITÀ DEI PROBLEMI DI USABILITÀ PER L'APPLICAZIONE DI NOTABILITY.

Dark Patterns - Notability

FORCED CONTINUITY

*“Tecnica manipolativa in cui un servizio richiede informazioni di pagamento per un **periodo di prova gratuito** per poi **attivare automaticamente l'abbonamento a pagamento senza avvisare** chiaramente l'utente”*

Notabilty come GoodNotes, ha il rinnovo automatizzato e non notificato dell’abbonamento. Anche il passaggio dalla prova gratuita alla versione completa non è notificato e il pagamamento è automatico.



Frasi Rilevanti - Notability

Registrate durante la Valutazione Olistica degli Esperti di Dominio

"Basta.. L'ho selezionato e non posso più cambiarlo. **Non ho possibilità di decisione**"

(in riferimento alle impostazioni rapide per la creazione di un documento)

"**Questa cosa è abbastanza scomoda**, dovrebbe rimanere fissa!"

(in riferimento alla barra degli strumenti)

"Le icone sono **meno intuitive** dell'altra"

(in riferimento agli strumenti per la modifica di un documento)

"Vediamo che fa salva... Assolutamente niente [...] **Potresti anche dirmelo dove me l'hai salvato!**"

(frase registrata durante l'esplorazione delle opzioni disponibili una volta selezionato un elemento con lo strumento "Lazo")

VALUTAZIONE EURISTICA:

Considerazioni Finali

PROBLEMI RILEVATI

Dalle valutazioni olistiche svolte dai valutatori, l'analisi di **GoodNotes** ha rilevato **21 problemi**, di cui solamente 1 con priorità alta.

Nel caso di **Notability** sono invece stati individuati **17 problemi**, di cui 3 con priorità alta.

EURISTICHE VIOLATE

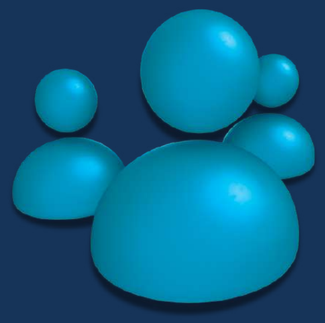
È stato osservato una frequenza di violazione delle euristiche quasi identica tra le due applicazioni, con **prevalenza** di violazioni delle **euristiche di cognizione** in entrambe.

DARK PATTERNS

Entrambe le applicazioni presentano il dark pattern ***Forced Continuity***. **GoodNotes** in più fa uso anche di ***Disguised Ads***, ***Price Comparison Prevention*** e ***Privacy Zuckering***.

TEST UTENTE

Task



Utenti Coinvolti

Efficacia

Efficienza

Espressioni Facciali

Conclusioni

Overview

Dopo la valutazione qualitativa, fatta tramite la Valutazione Euristica, vogliamo valutare le due applicazioni anche da un punto di vista **quantitativo**.

L'obiettivo in particolare è di confrontare **l'efficacia** e **l'efficienza** dei due sistemi.
Per fare ciò abbiamo sottoposto un ulteriore campione di utenti a **3 task** da svolgere su ciascuna applicazione.

A seguito di questo processo, sono stati effettuati **test statistici** sui dati raccolti in modo da verificare la **significatività statistica** delle affermazioni.

EFFICACIA

= grado di correttezza, accuratezza, completezza con cui l'utente raggiunge i suoi obiettivi.

EFFICIENZA

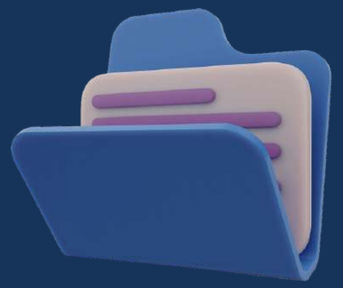
= rapporto tra le risorse usate per raggiungere il risultato e quelle effettivamente necessarie.

Task

I task sono stati concepiti in modo che fossero realistici, eseguibili e non troppo specifici.

I tre task presentati agli utenti sono gli stessi sia per GoodNotes che per Notability e sono stati pensati per testare le principali funzionalità delle applicazioni: creazione di **cartelle e documenti**, utilizzo di **strumenti** per prendere appunti (anche tramite l'utilizzo di Apple Pencil) e l'impiego di funzionalità per sfruttare al meglio **documenti importati**.

I task sono di seguito riportati e verranno presentati più nel dettaglio nelle slide seguenti.



TASK 1
Crea cartelle e documenti

TASK 2
Usa gli strumenti

TASK 3
Cerca una parola nel documento

Task 1

CREA CARTELLE E DOCUMENTI

“Usa l’applicazione per creare una cartella in un punto qualsiasi, crea un documento nella cartella appena creata, disegna 3 fiori di colore diverso sul nuovo documento.”

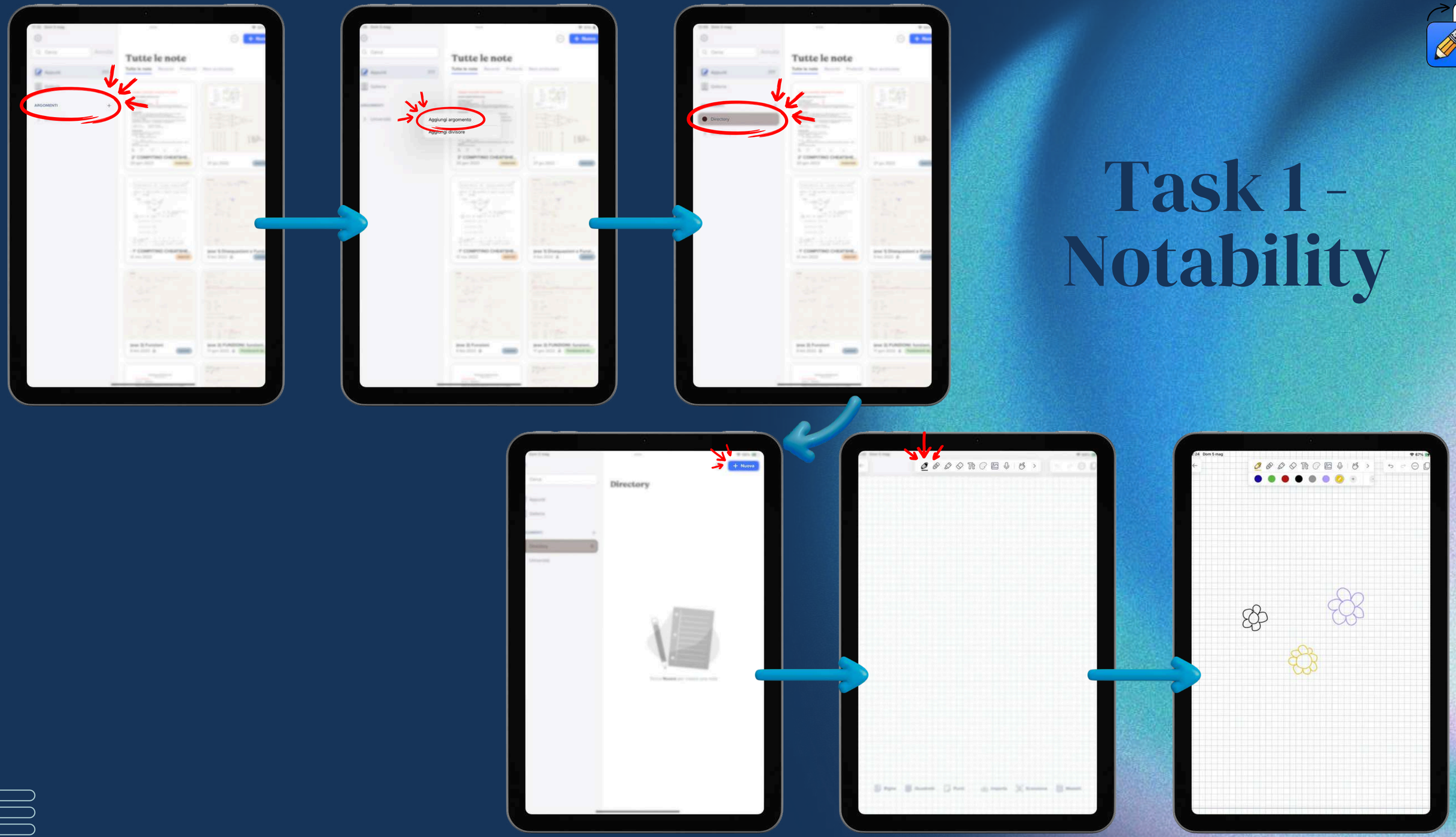
STEP 1
Crea una cartella

STEP 2
Crea un documento nella cartella

STEP 3
Disegna 3 fiori di colore diverso

Task 1 - GoodNotes

Task 1 - Notability



Task 2

USA GLI STRUMENTI

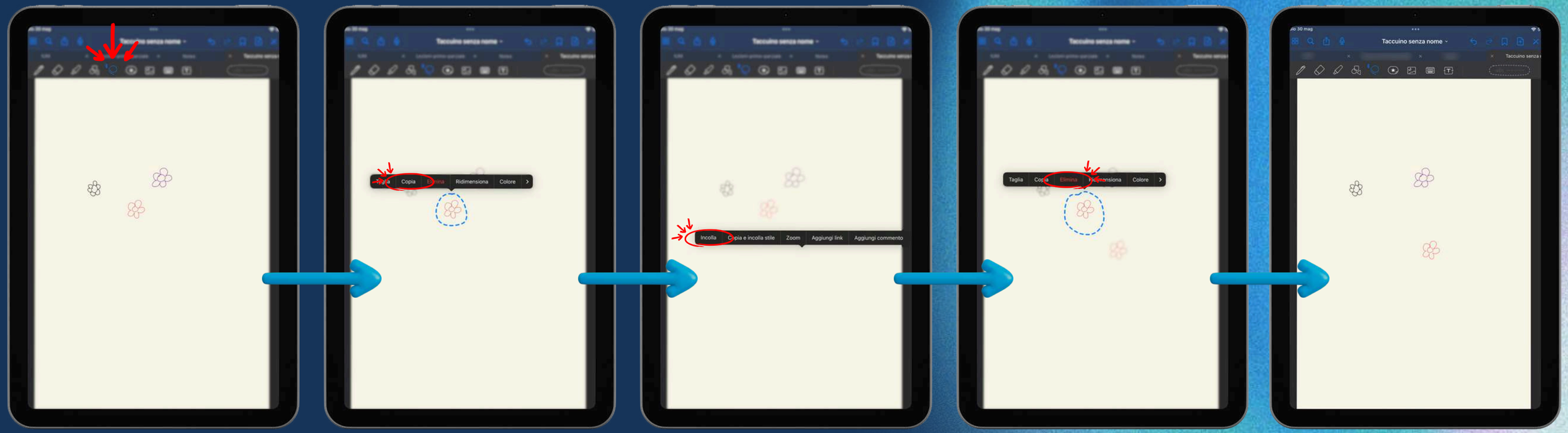
“Seleziona uno qualsiasi dei fiori che hai creato (nel Task 1), creane una copia ed elimina la versione originale.”

STEP 1
Seleziona uno dei fiori

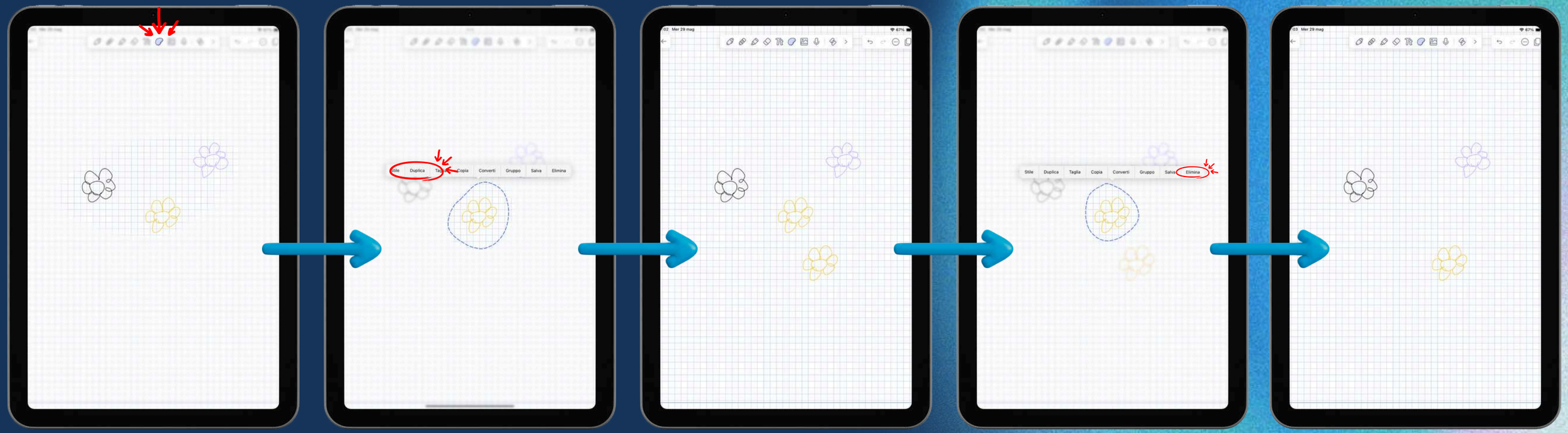
STEP 2
Crea una copia

STEP 3
Elimina l’originale

Task 2 - GoodNotes



Task 2 - Notability



Task 3

CERCA UNA PAROLA IN UN DOCUMENTO

“Apri un documento qualsiasi, cerca se esiste la parola ”prova” nel documento e se esiste evidenziala e sottolineala.”

STEP 1
Apri un documento

STEP 2
Cerca la parola “prova”

STEP 3
Evidenzia e sottolinea

Task 3 - GoodNotes

Task 3 - Notability

Utenti coinvolti

È stato coinvolto un campione formato da **16 utenti**, compresi i 3 esperti di dominio che avevano già preso parte alla valutazione euristica.

ID UTENTE	GENERE	FASCIA D'ETÀ	OCCUPAZIONE	FAMILIARITÀ
ED1	M	<25	Studente	Alta
ED2	M	<25	Studente	Media
ED3	F	<25	Studente	Media
U4	F	<25	Studente	Bassa
U5	F	<25	Studente	Alta
U6	M	>50	Lavoratore	Bassa
U7	M	<25	Studente	Alta
U8	F	tra 25 e 50	Lavoratore	Bassa
U9	M	<25	Studente	Media
U10	M	<25	Studente	Media
U11	M	<25	Studente	Alta
U12	F	<25	Studente	Media
U13	M	tra 25 e 50	Studente	Alta
U14	F	<25	Studente	Media
U15	F	<25	Lavoratore	Bassa
U16	F	<25	Studente	Bassa

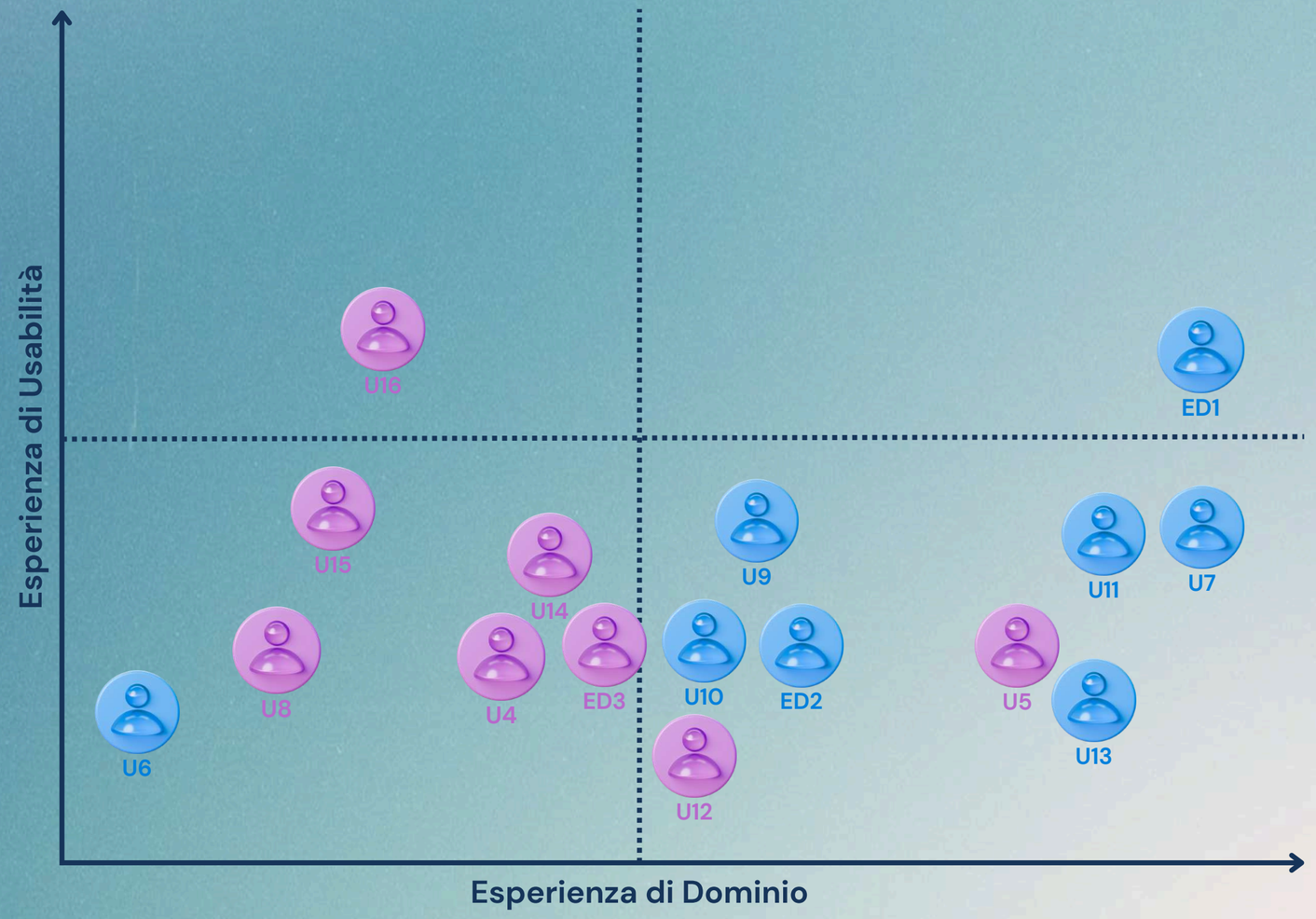
Matrice di Expertise

La matrice di expertise permette di vedere

- l'**esperienza di usabilità**
- l'**esperienza di dominio**

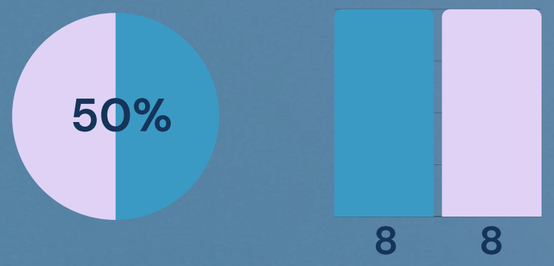
di ciascun utente coinvolto.

In particolare nel nostro caso si può osservare una conoscenza dell'**usabilità medio-bassa** e una conoscenza più **equamente distribuita** del **dominio**.

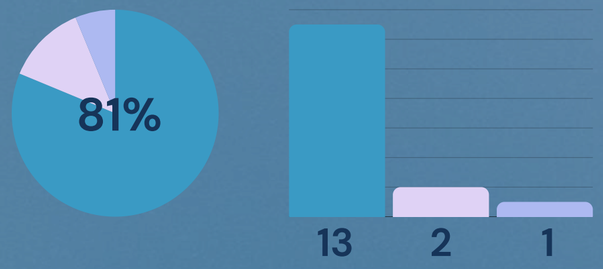


Composizione del campione

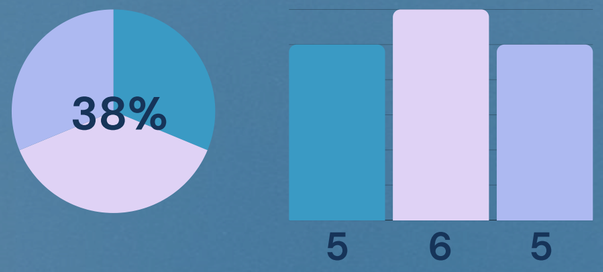
GENERE
- Maschio – Femmina



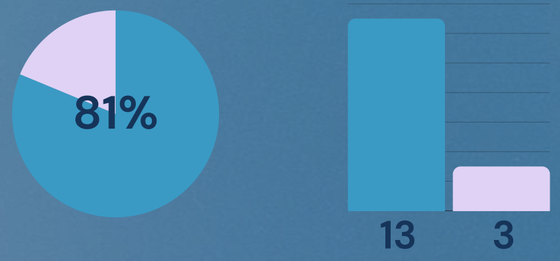
ETÀ
- <25



FAMILIARITÀ
- Media



OCCUPAZIONE
- Studente



I valutatori coinvolti sono:
- distribuiti equamente per quanto riguarda il genere e la familiarità
- prevalentemente di età **inferiore ai 25 anni**
- per la maggior parte **studenti**.

Efficacia

Per confrontare l'efficacia delle due applicazioni saranno analizzati i dati raccolti durante i Test Utente. L'efficacia è stata suddivisa in:

EFFICACIA RELATIVA:
Sono stati verificati gli esiti dei task, ovvero se gli utenti fossero riusciti a completare in modo corretto i compiti assegnati. Tali esiti sono stati classificati come:

- **successo**
- **successo parziale*** (esecuzione non ottimale o con suggerimenti)
- **insuccesso**

Per poi confrontare i tassi di errori medi per ogni task, calcolati come il rapporto tra il numero di errori e il numero totale dei tentativi.

EFFICACIA ASSOLUTA:
Viene richiesto che il tasso di errore per ogni task non superi il tasso ottimale, scelto al 5% (ovvero un errore* nel nostro caso).

*nella nostra analisi, per il calcolo del tasso di errore medio i successi parziali sono stati contati come insuccessi e il tasso ottimale è stato arrotondato per eccesso da 0,8 ad 1 errore.

Per verificare la significatività statistica dei dati raccolti sono stati utilizzati diversi test statistici. Per l'analisi dell'efficacia è stato utilizzato in particolare il **Test esatto di Fisher** con un **intervallo di confidenza** del **95%**; ciò significa che un **p-value inferiore al 5%** indica la presenza di una **differenza statisticamente significativa** tra le due applicazioni.

Ipotesi nulla (H_0): "Non ci sono differenze statisticamente significative tra l'applicazione di GoodNotes e l'applicazione di Notability"

Efficacia - Matrice degli errori

	TASK 1		TASK 2		TASK3	
	GoodNotes	Notability	GoodNotes	Notability	GoodNotes	Notability
ED1						
ED2						
ED3						
U4						
U5						
U6						
U7						
U8						
U9						
U10						
U11						
U12						
U13						
U14						
U15						
U16						

SUCCESSO SUCCESSO PARZIALE INSUCCESSO

Considerando i **successi parziali** come **insuccessi**, i tassi di errore risultanti sono i seguenti:

TASSO DI ERRORE GOODNOTES:
25%

TASSO DI ERRORE NOTABILITY:
41,7%

Efficacia - Task 1

È stato effettuato un **Test esatto di Fisher** per confrontare il rapporto tra successi ed errori (successi parziali e insuccessi) nel compiere il **Task 1** usando GoodNotes e Notability.

In base a tale test **NON è stata trovata una differenza statisticamente significativa** riguardo al numero di successi:

- *GoodNotes*: 13/16
- *Notability*: 10/16

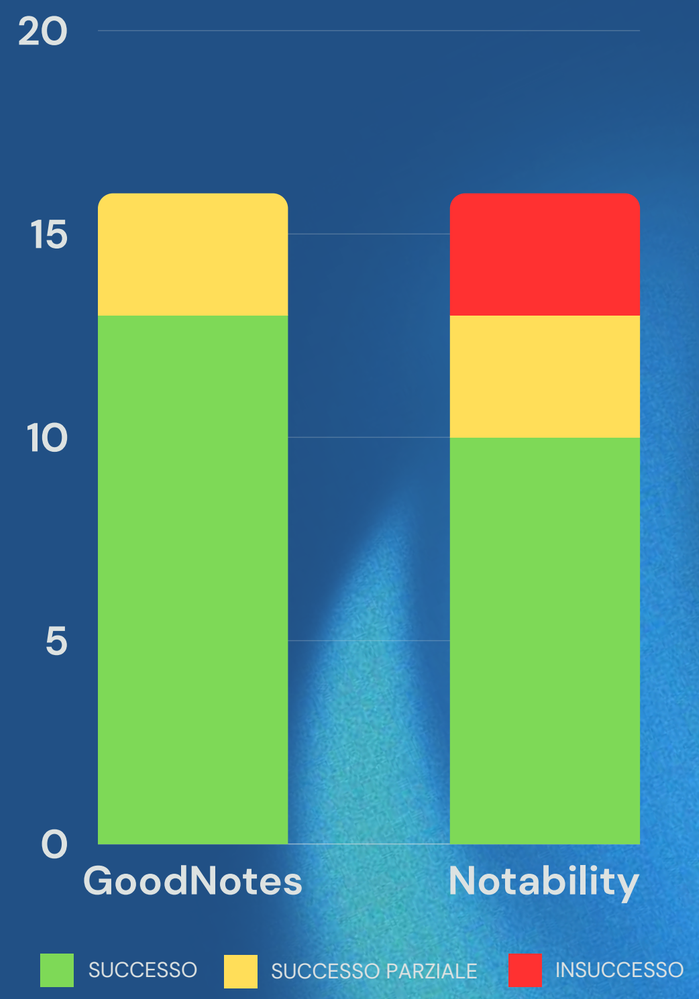
e al numero di errori*:

- *GoodNotes*: 3/16
- *Notability*: 6/16

p-value: 0.4331

Questo risultato suggerisce che gli utenti **NON commettono un numero significativamente diverso di errori** nell'eseguire il Task 1 utilizzando GoodNotes piuttosto che Notability.

*i successi parziali sono contati come insuccessi



Efficacia - Task 2

È stato effettuato un **Test esatto di Fisher** per confrontare il rapporto tra successi ed errori (successi parziali e insuccessi) nel compiere il **Task 2** usando GoodNotes e Notability.

In base a tale test **NON è stata trovata una differenza statisticamente significativa** riguardo al numero di successi:

- *GoodNotes*: 12/16
- *Notability*: 15/16

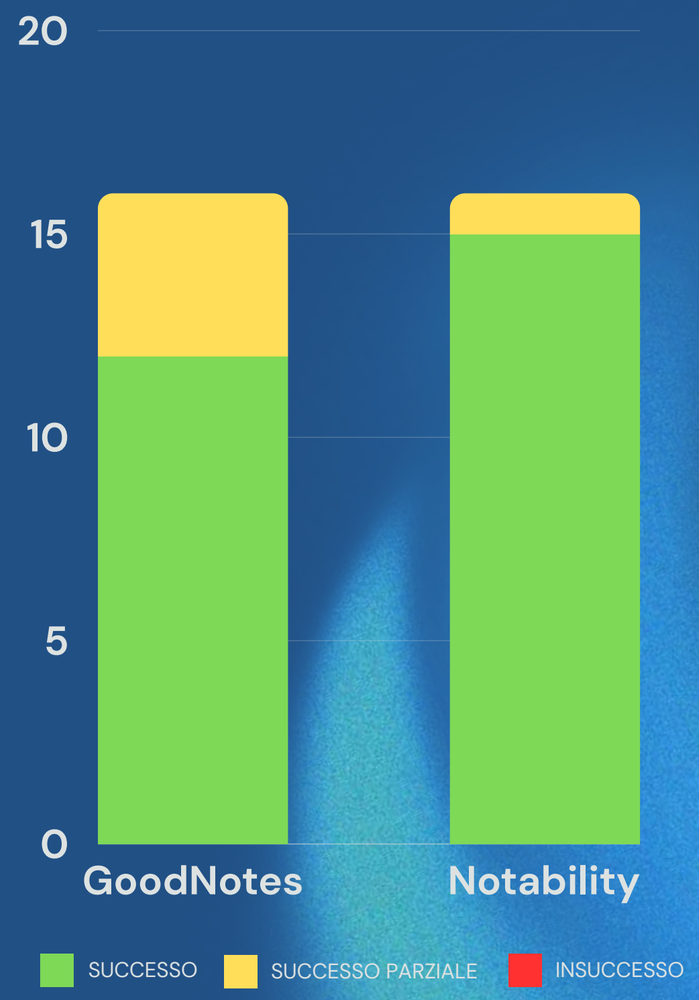
e al numero di errori*:

- *GoodNotes*: 4/16
- *Notability*: 1/16

p-value: 0.3326

Questo risultato suggerisce che gli utenti **NON commettono un numero significativamente diverso di errori** nell'eseguire il Task 2 utilizzando GoodNotes piuttosto che Notability.

*i successi parziali sono contati come insuccessi



Efficacia - Task 3

È stato effettuato un **Test esatto di Fisher** per confrontare il rapporto tra successi ed errori (successi parziali e insuccessi) nel compiere il **Task 3** usando GoodNotes e Notability.

In base a tale test **è stata trovata una differenza statisticamente significativa** riguardo al numero di successi:

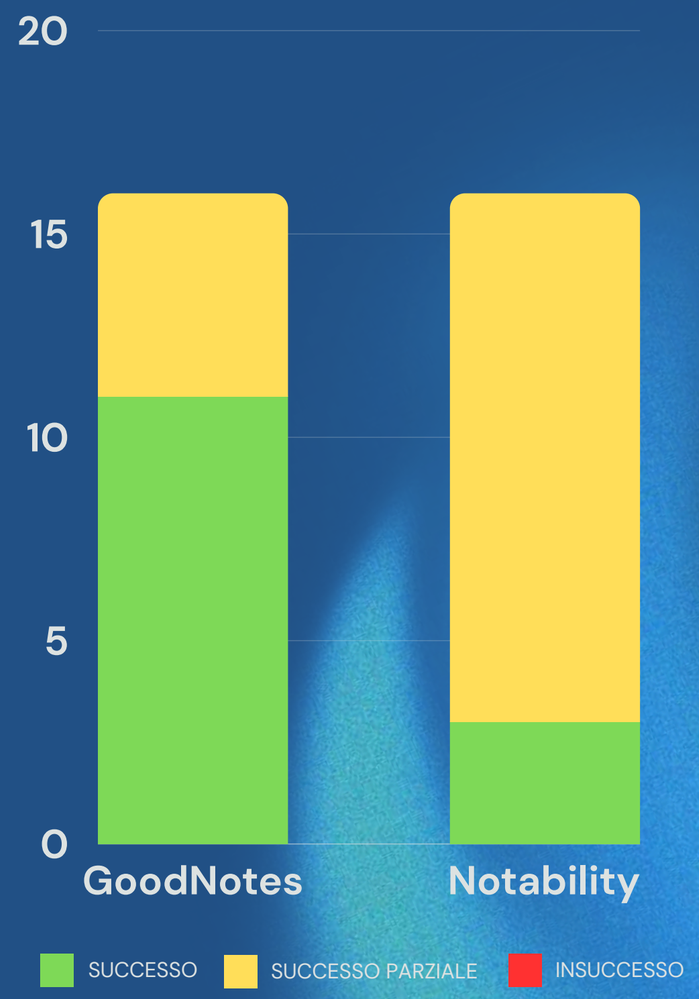
- *GoodNotes*: 11/16
- *Notability*: 3/16

e al numero di errori*:

- *GoodNotes*: 5/16
- *Notability*: 13/16

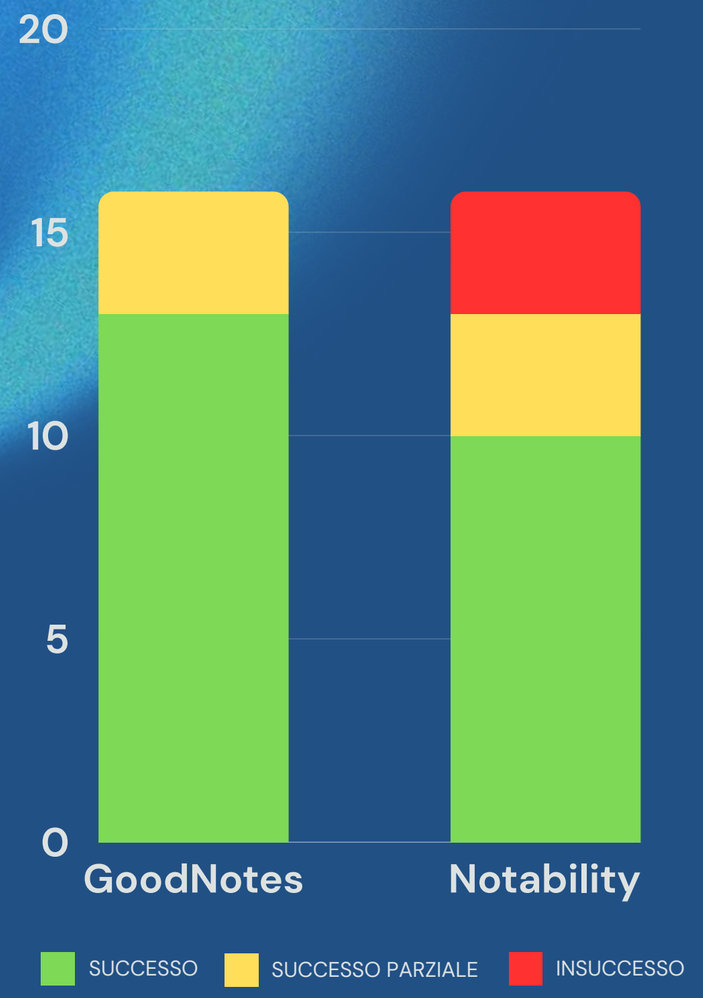
p-value: 0.0113

Questo risultato suggerisce che gli utenti **commettono un numero significativamente maggiore di errori** nell'eseguire il Task 3 utilizzando l'applicazione di Notability rispetto a quella di GoodNotes.



*i successi parziali sono contati come insuccessi

Efficacia assoluta - Task 1



Effettuando il Test esatto di Fisher **NON è stata evidenziata alcuna differenza statisticamente significativa** confrontando

- *numero di errori* GoodNotes*: 3/16
- *numero di errori* Notability*: 6/16

con **soglia ottimale tasso di errore**: 1/16

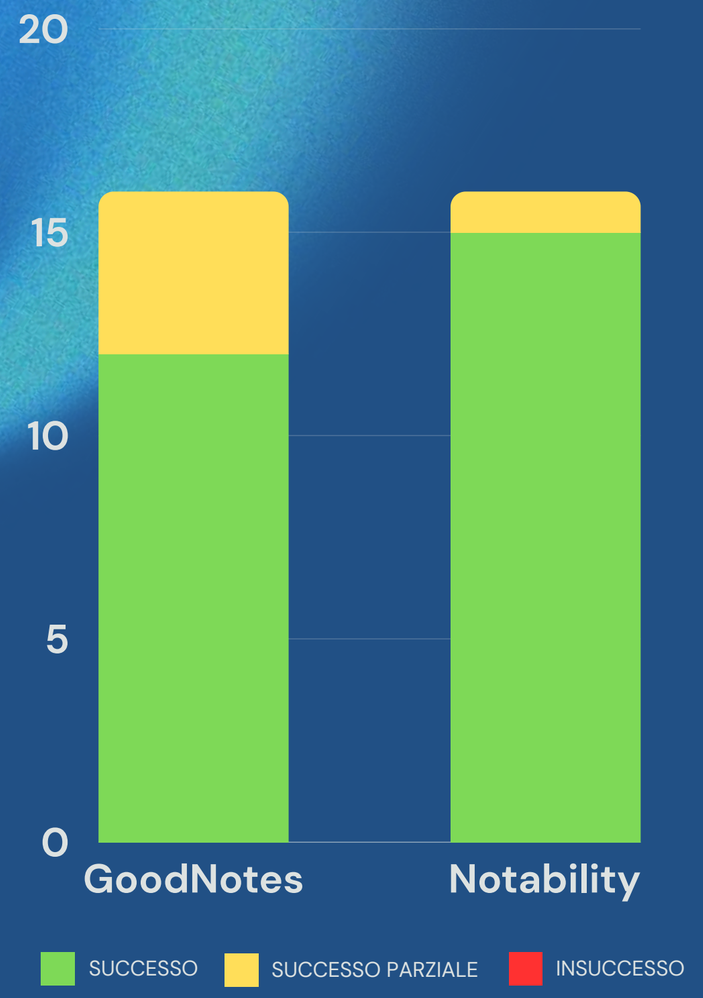
p-value GoodNotes: 0.5996

p-value Notability: 0.0829

Questo risultato suggerisce che gli utenti **NON commettono un numero significativamente maggiore di errori** rispetto alla soglia ottimale nell'eseguire il Task 1 utilizzando GoodNotes piuttosto che Notability.

*i successi parziali sono contati come insuccessi

Efficacia assoluta - Task 2



Effettuando il Test esatto di Fisher **NON è stata evidenziata alcuna differenza statisticamente significativa** confrontando

- *numero di errori* GoodNotes*: 4/16
- *numero di errori* Notability*: 1/16

con **soglia ottimale tasso di errore**: 1/16

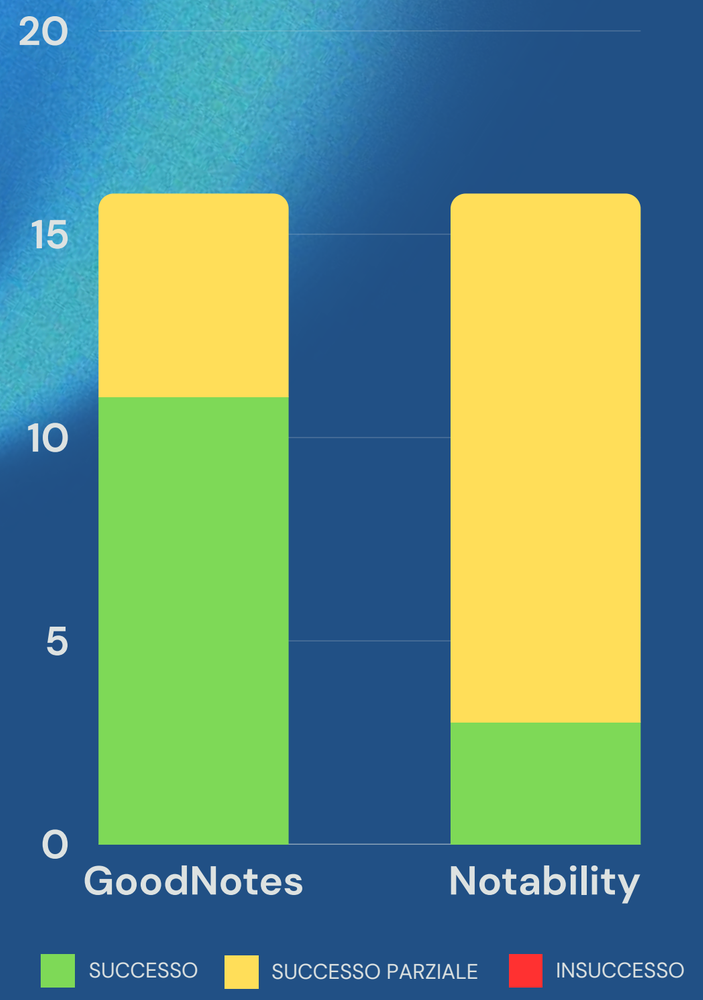
p-value GoodNotes: 0.3326

p-value Notability: 1

Questo risultato suggerisce che gli utenti **NON commettono un numero significativamente maggiore di errori** rispetto alla soglia ottimale nell'eseguire il Task 2 utilizzando GoodNotes piuttosto che Notability.

*i successi parziali sono contati come insuccessi

Efficacia assoluta - Task 3



Effettuando il Test esatto di Fisher **è stata evidenziata una forte differenza statisticamente significativa** confrontando

- *numero di errori* GoodNotes*: 5/16
- *numero di errori* Notability*: 13/16

con **soglia ottimale tasso di errore**: 1/16

p-value GoodNotes: 0.1719

p-value Notability: <0.0001

Questo risultato suggerisce che gli utenti **commettono un numero significativamente maggiore di errori** rispetto alla soglia ottimale nell'eseguire il Task 3 utilizzando Notability piuttosto che GoodNotes.

*i successi parziali sono contati come insuccessi

Efficienza

Per confrontare l'efficienza delle due applicazioni saranno analizzati i dati raccolti durante i Test Utente.
L'efficienza è stata suddivisa in:

EFFICIENZA RELATIVA:
I **tempi di esecuzione dei task** di ciascun utente (misurati riguardando i video registrati e reperibili nell'appendice) verranno analizzati e confrontati con i **tempi di esecuzione medi** per ogni task.

EFFICIENZA ASSOLUTA:
Le performance degli utenti vengono confrontate con l'**OET** (Optimal Execution Time), cioè una stima del tempo di esecuzione ottimale per quel task

Per verificare la significatività statistica dei dati raccolti sono stati utilizzati diversi test statistici.
Per l'analisi dell'efficienza è stato utilizzato in particolare il **test t di Student per campioni accoppiati** con un **intervallo di confidenza del 95%**; ciò significa che un **p-value inferiore al 5%** indica la presenza di una **differenza statisticamente significativa** tra le due applicazioni.

Ipotesi nulla (H_0): "Non ci sono differenze statisticamente significative tra l'applicazione di GoodNotes e l'applicazione di Notability"

Efficienza

Come precedentemente accennato, grazie alle registrazioni di ciascun task per ogni utente è stato possibile registrare i **tempi di esecuzione**, riportati qui a fianco.

Dalla media si può osservare che per i **task 1 e 3** gli utenti hanno impiegato **meno tempo** utilizzando **GoodNotes**, mentre per il **task 2 Notability** ha registrato dei **tempi di esecuzione più rapidi**.

	TASK 1		TASK 2		TASK3	
	GoodNotes	Notability	GoodNotes	Notability	GoodNotes	Notability
ED1	0:55	0:45	0:30	0:18	0:32	0:42
ED2	1:00	1:22	0:22	0:24	1:19	1:26
ED3	0:51	1:53	0:14	0:19	0:33	1:02
U4	1:16	1:17	1:10	0:40	0:53	0:42
U5	0:58	0:40	0:45	0:30	2:13	2:22
U6	3:34	2:30	1:31	1:20	0:55	0:51
U7	0:41	2:10	0:18	0:20	2:00	3:55
U8	1:59	1:28	1:59	0:42	0:58	1:28
U9	0:43	0:56	0:12	0:15	0:26	1:17
U10	1:14	1:09	0:50	0:34	1:02	2:22
U11	0:50	1:09	0:22	0:32	0:48	1:54
U12	1:01	1:32	0:23	0:36	1:04	1:15
U13	1:29	1:08	0:29	0:23	1:17	1:20
U14	1:22	2:19	0:55	0:52	1:27	1:59
U15	1:39	0:40	1:30	0:46	2:32	3:43
U16	0:42	0:45	0:25	0:32	0:23	1:23
MEDIA	*1:15*	*1:21*	*0:44*	*0:34*	*1:09*	*1:43*
DEV. ST.	*0:43*	*0:35*	0:32	*0:16*	*0:37*	*0:58*

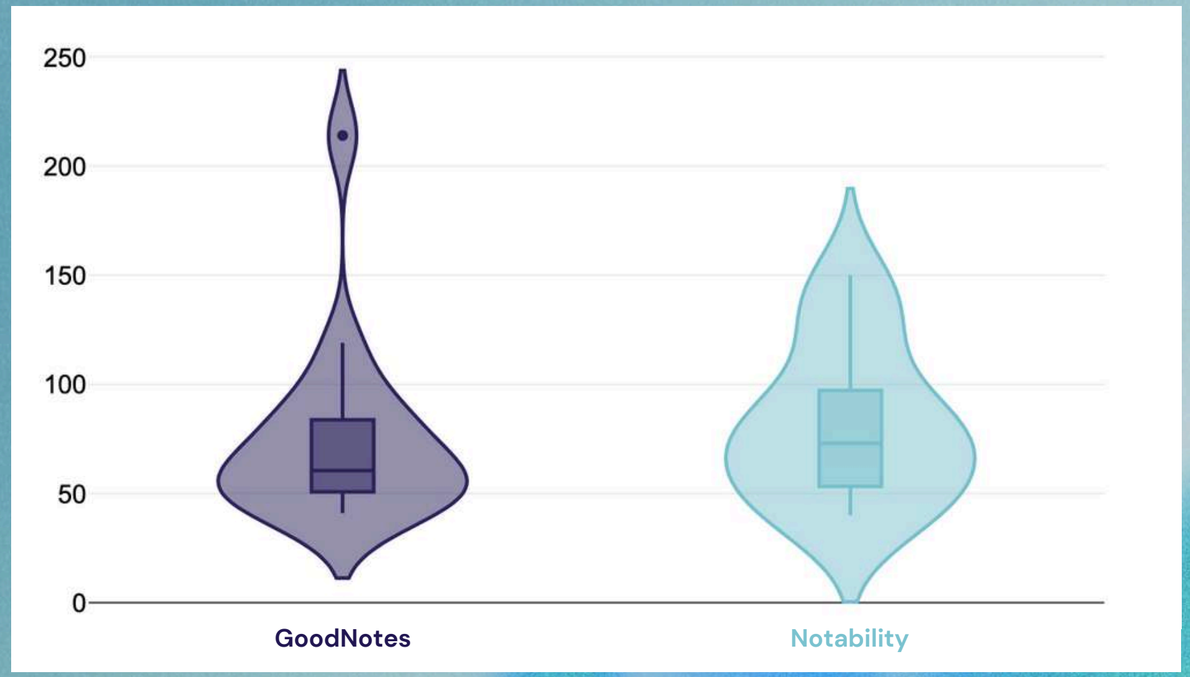
Efficienza - Task 1

Svolgendo un test t di Student per campioni accoppiati sui dati raccolti dai 16 utenti **NON si è potuta osservare alcuna differenza statisticamente significativa** confrontando i tempi di esecuzione:

- tempo medio GoodNotes: 1:15
- tempo medio Notability: 1:21

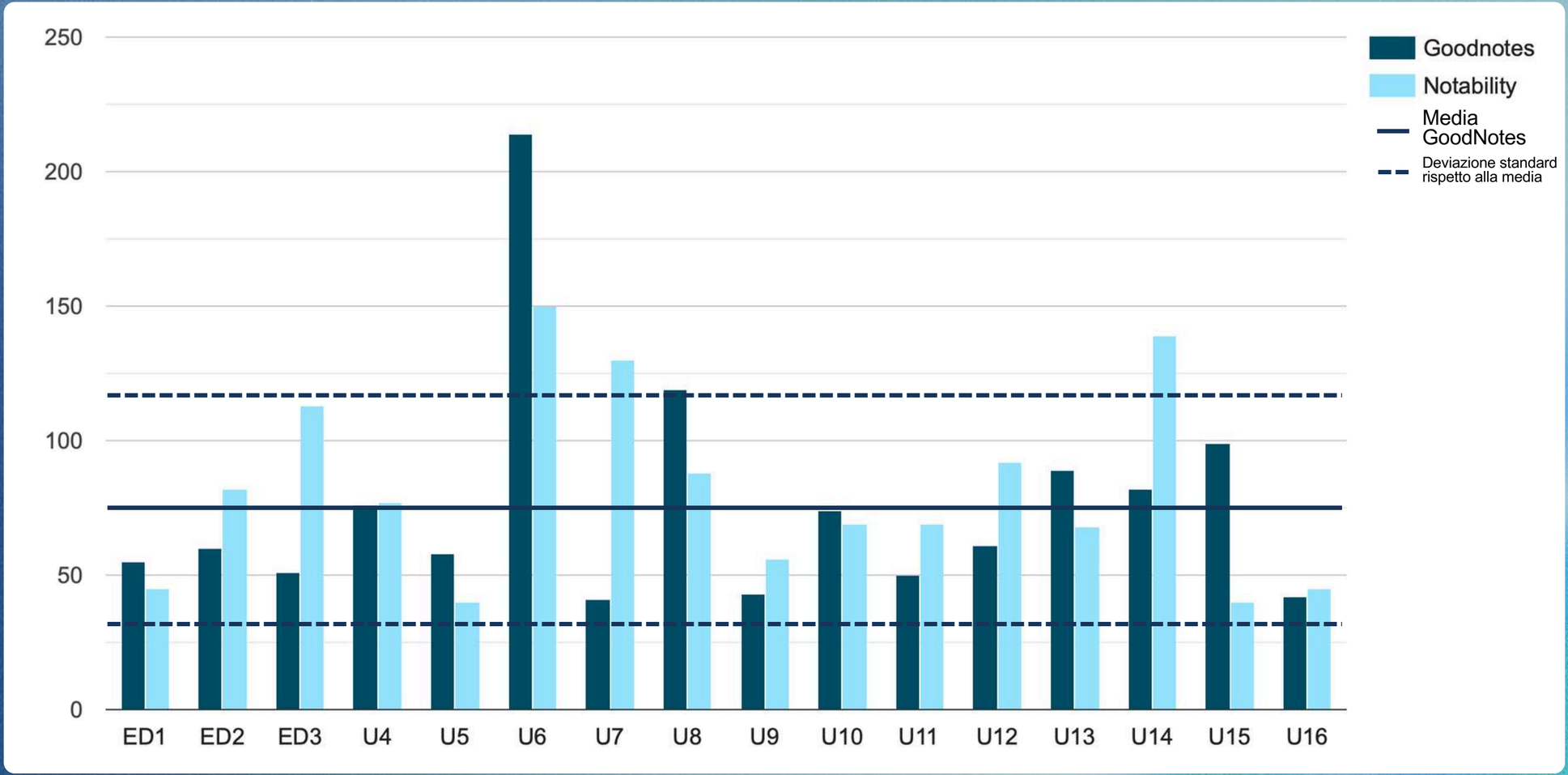
t = 0.5351
p = 0.6004

Questo risultato suggerisce che **gli utenti NON impiegano un tempo significativamente diverso** per eseguire il Task 1 con GoodNotes piuttosto che con Notability.

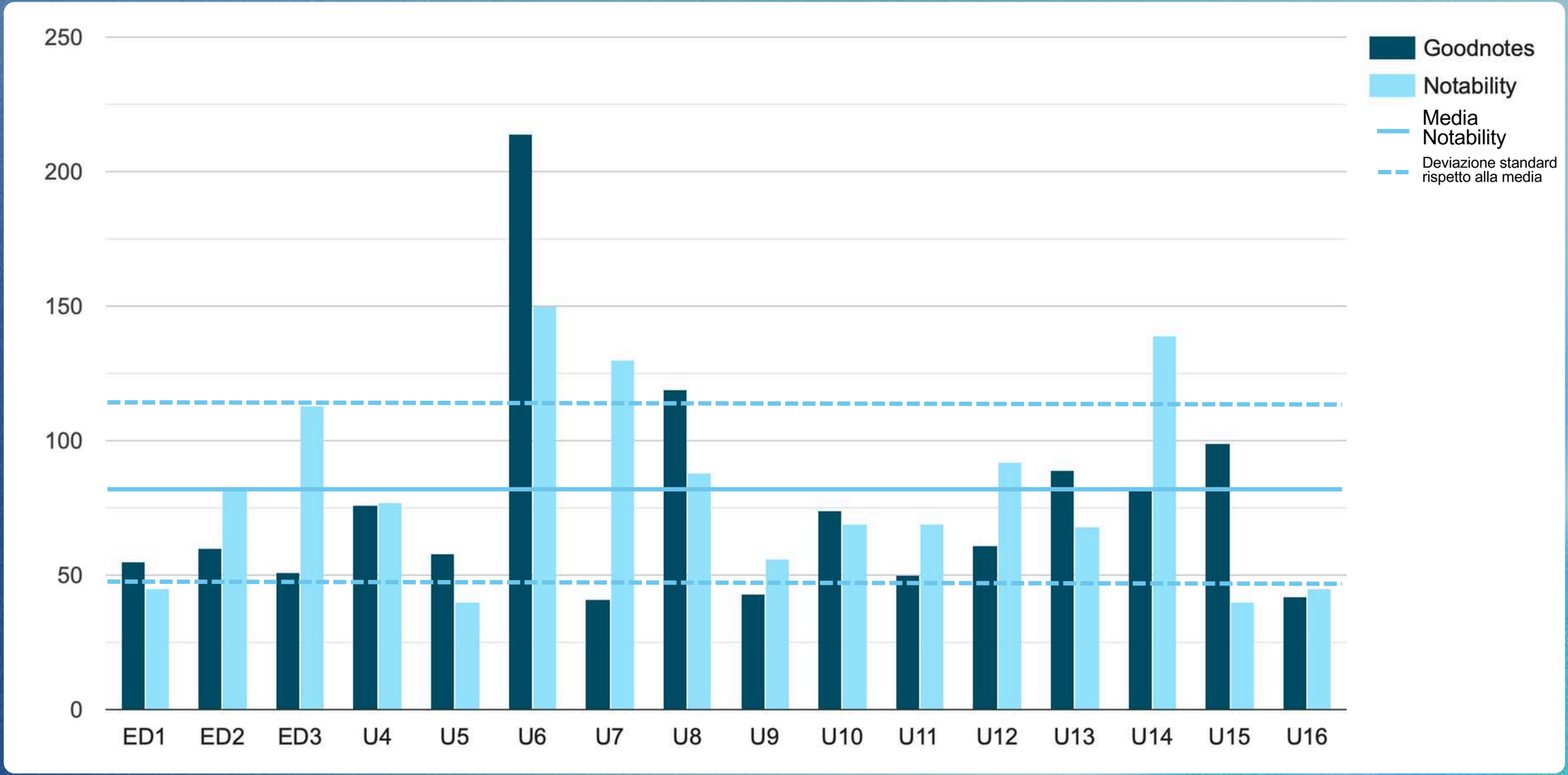


APPLICAZIONE	Baffo Inferiore	Primo Quartile	Mediana	Terzo Quartile	Baffo Superiore	Media
GoodNotes	0:41	0:50	1:00	1:23	1:59	1:15
Notability	0:40	0:53	1:13	1:37	2:30	1:21

Efficienza - tempi di esecuzione Task 1 (GoodNotes)



Efficienza - tempi di esecuzione Task 1 (Notability)



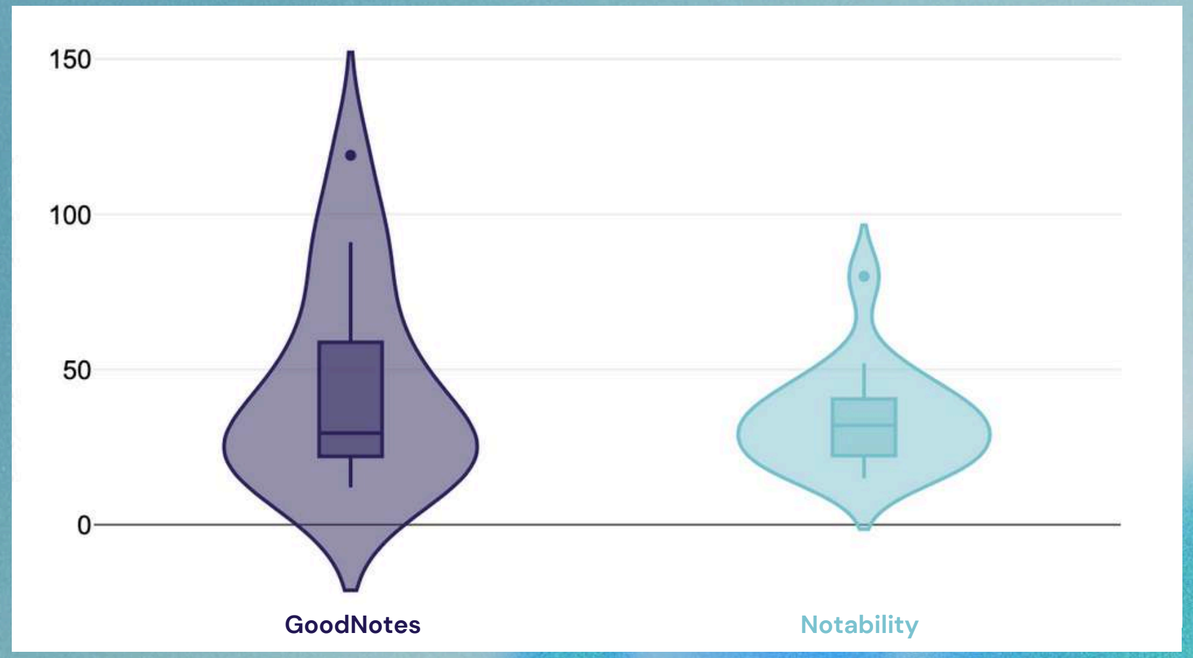
Efficienza - Task 2

Svolgendo un test t di Student per campioni accoppiati sui dati raccolti dai 16 utenti **NON si è potuta osservare alcuna differenza statisticamente significativa** confrontando i tempi di esecuzione:

- tempo medio GoodNotes: 0:44
- tempo medio Notability: 0:34

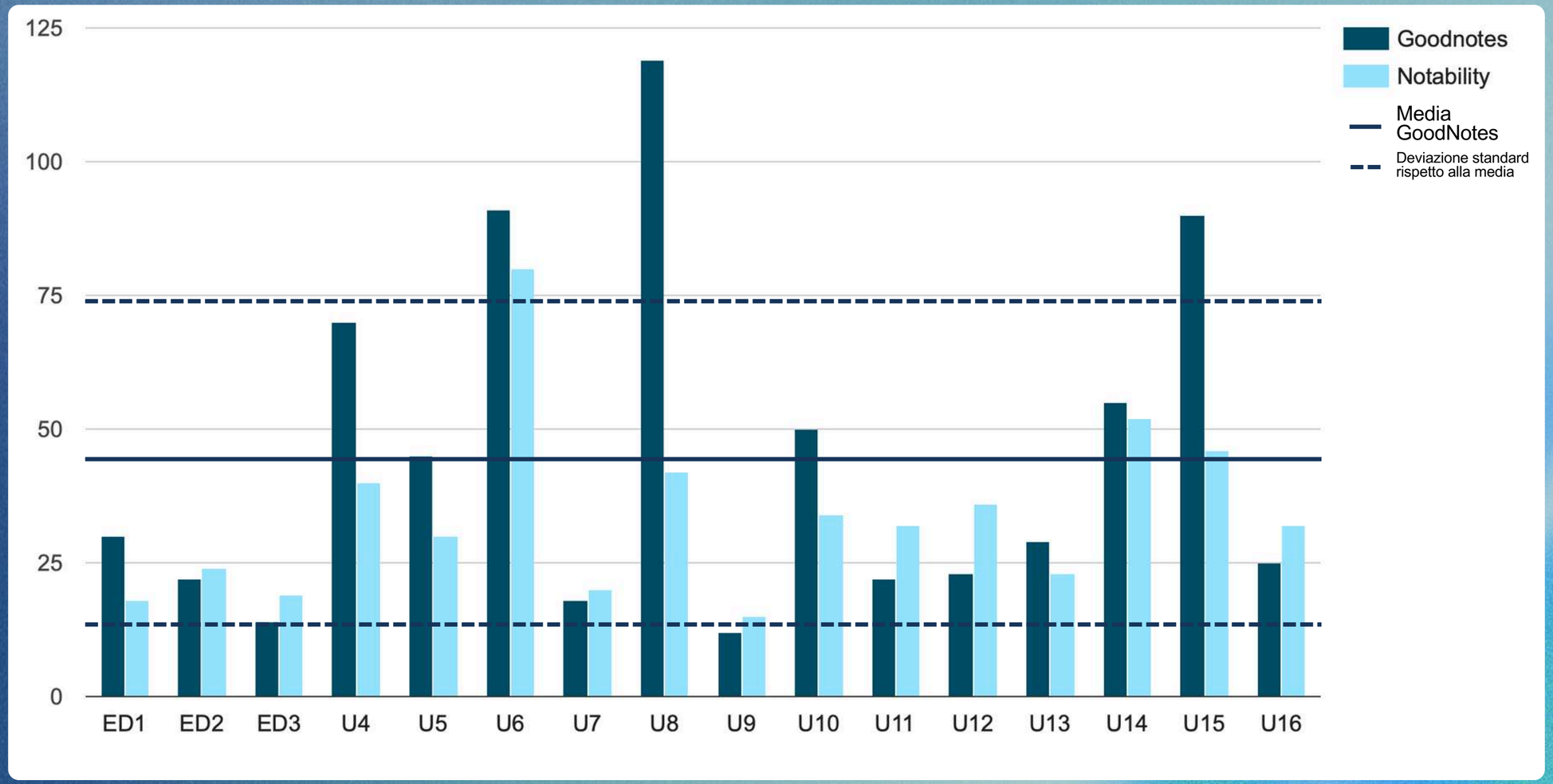
t = 1.8542
p = 0.0835

Questo risultato suggerisce che **gli utenti NON impiegano un tempo significativamente diverso** per eseguire il Task 2 con GoodNotes piuttosto che con Notability.

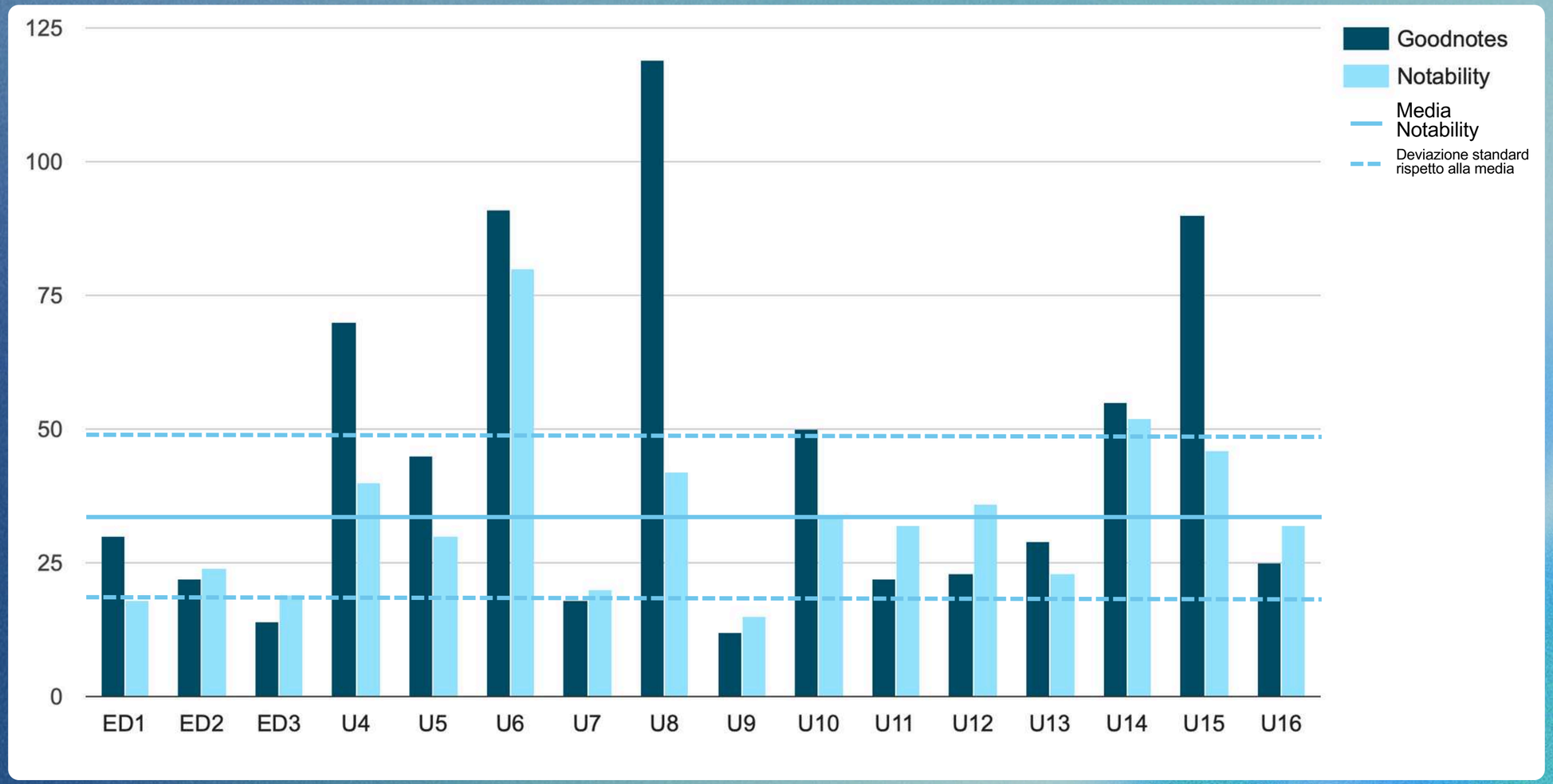


APPLICAZIONE	Baffo Inferiore	Primo Quartile	Mediana	Terzo Quartile	Baffo Superiore	Media
GoodNotes	0:12	0:22	0:29	0:59	1:31	0:44
Notability	0:15	0:22	0:32	0:40	0:52	0:34

Efficienza - tempi di esecuzione Task 2 (GoodNotes)



Efficienza - tempi di esecuzione Task 2 (Notability)



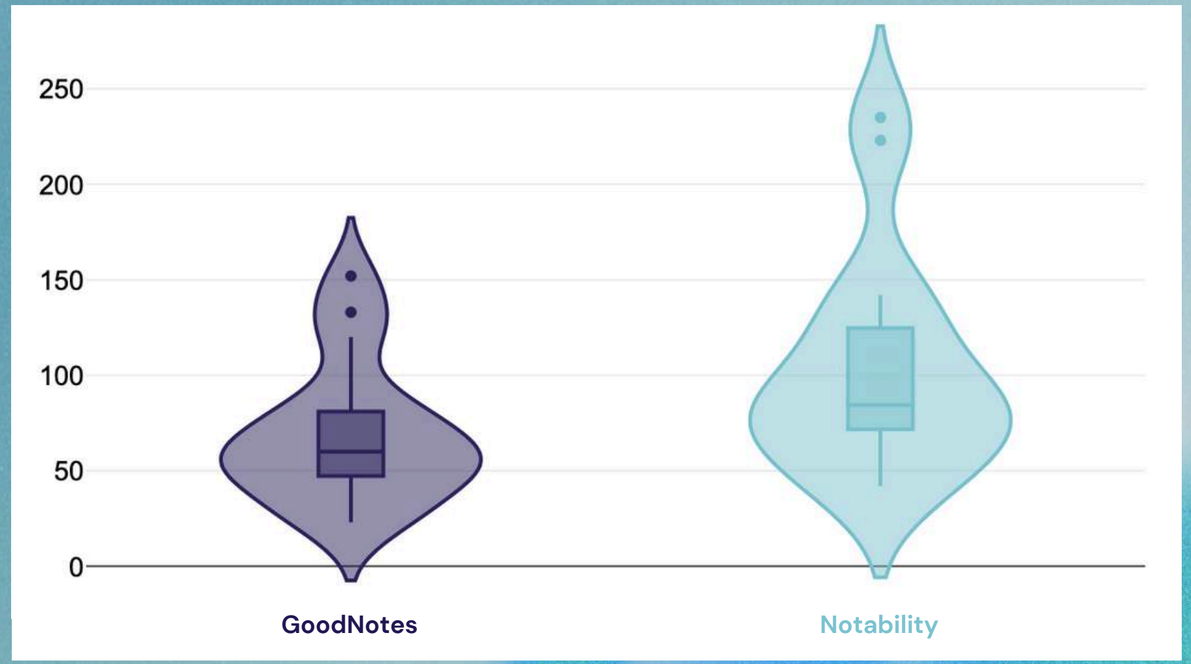
Efficienza - Task 3

Svolgendo un test t di Student per campioni accoppiati sui dati raccolti dai 16 utenti **è stata osservata una differenza statisticamente significativa** confrontando i tempi di esecuzione:

- tempo medio GoodNotes: 1:09
- tempo medio Notability: 1:44

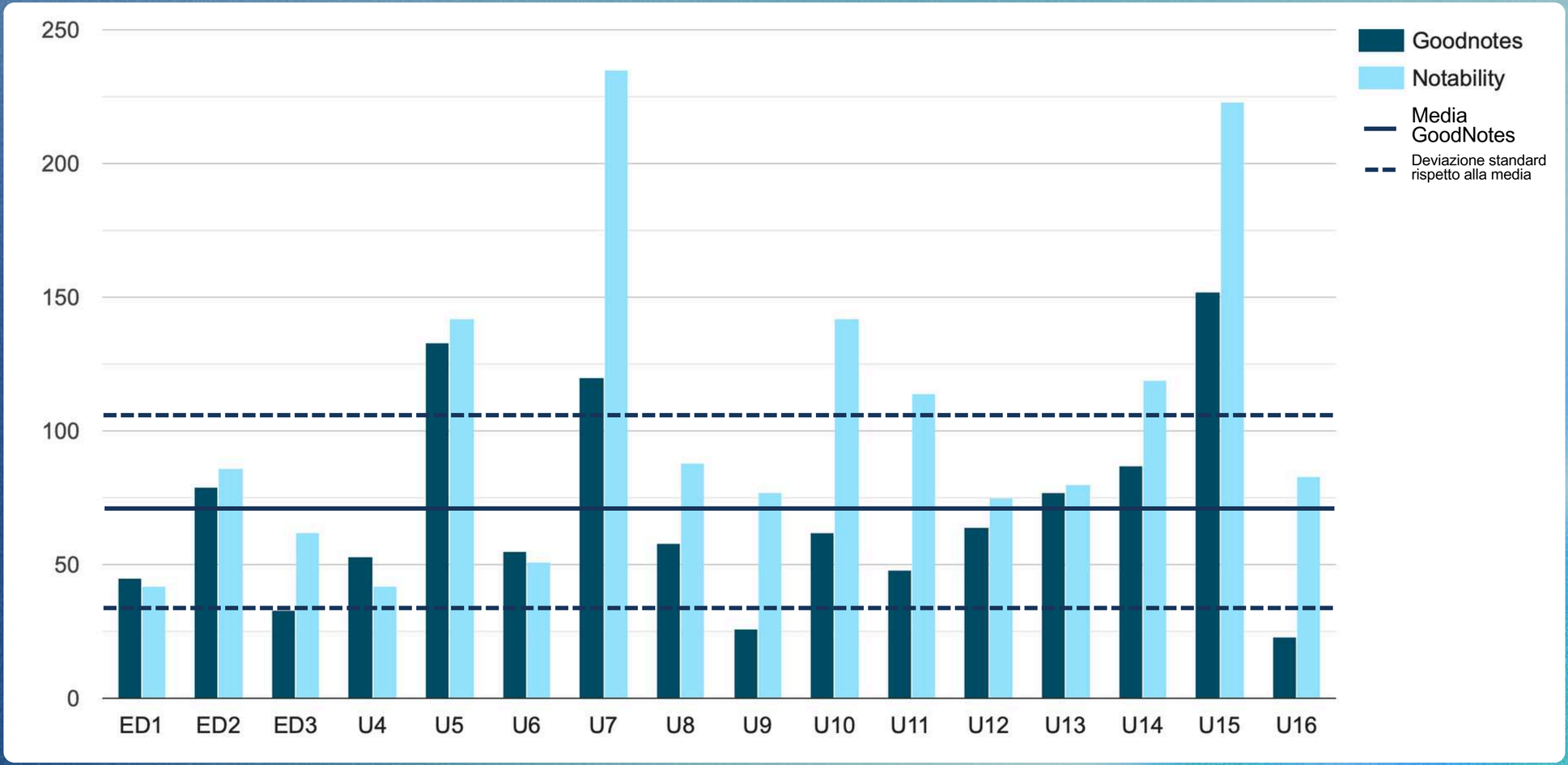
t = 3.7607
p = 0.0019

Questo risultato suggerisce che **gli utenti impiegano un tempo significativamente maggiore** per eseguire il Task 3 con Notability rispetto che con GoodNotes.

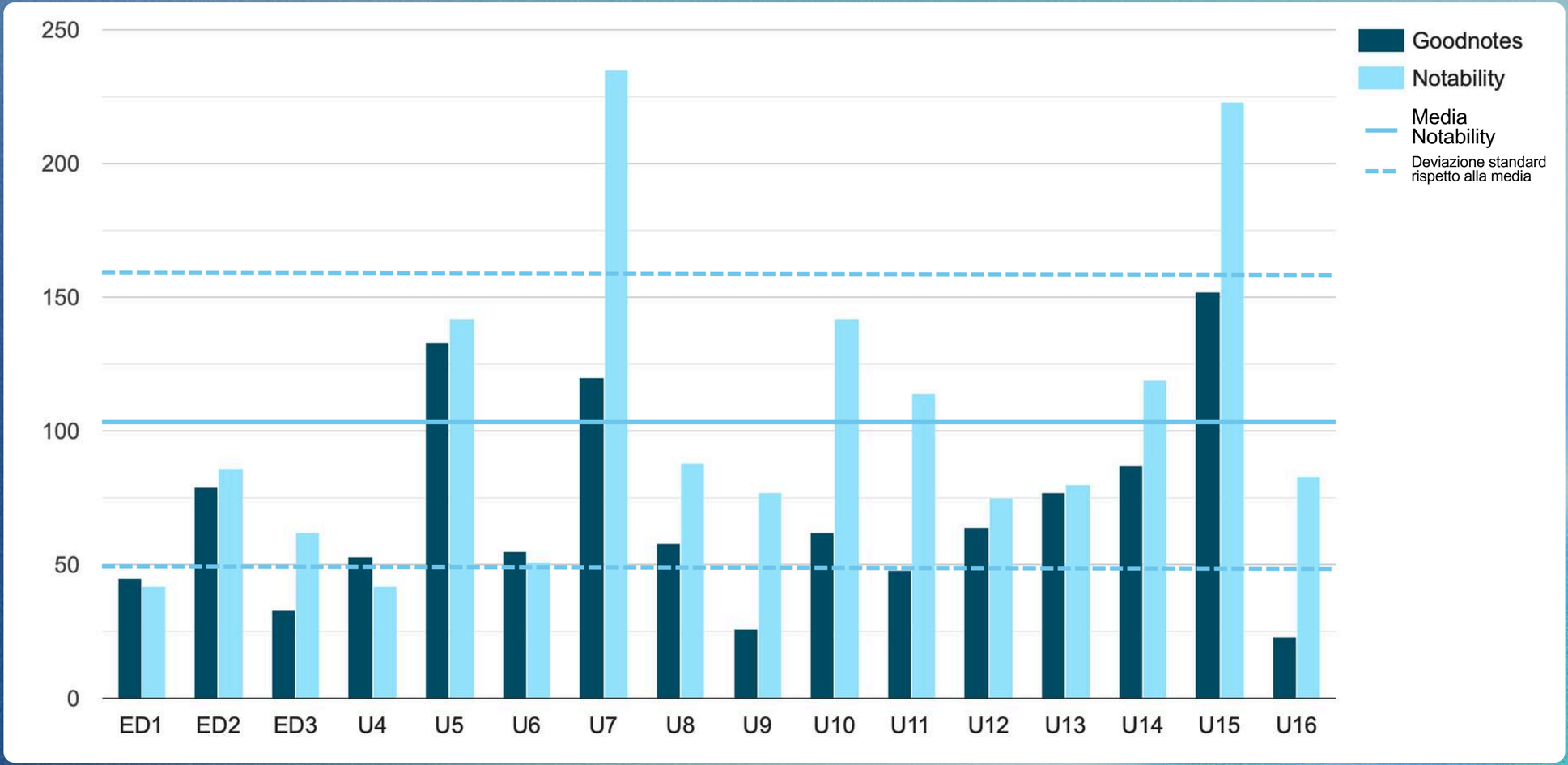


APPLICAZIONE	Baffo Inferiore	Primo Quartile	Mediana	Terzo Quartile	Baffo Superiore	Media
GoodNotes	0:23	0:47	1:00	1:21	2:00	1:09
Notability	0:42	1:12	1:24	2:05	2:22	1:44

Efficienza - tempi di esecuzione Task 3 (GoodNotes)



Efficienza - tempi di esecuzione Task 3 (Notability)



Efficienza assoluta - Task 1

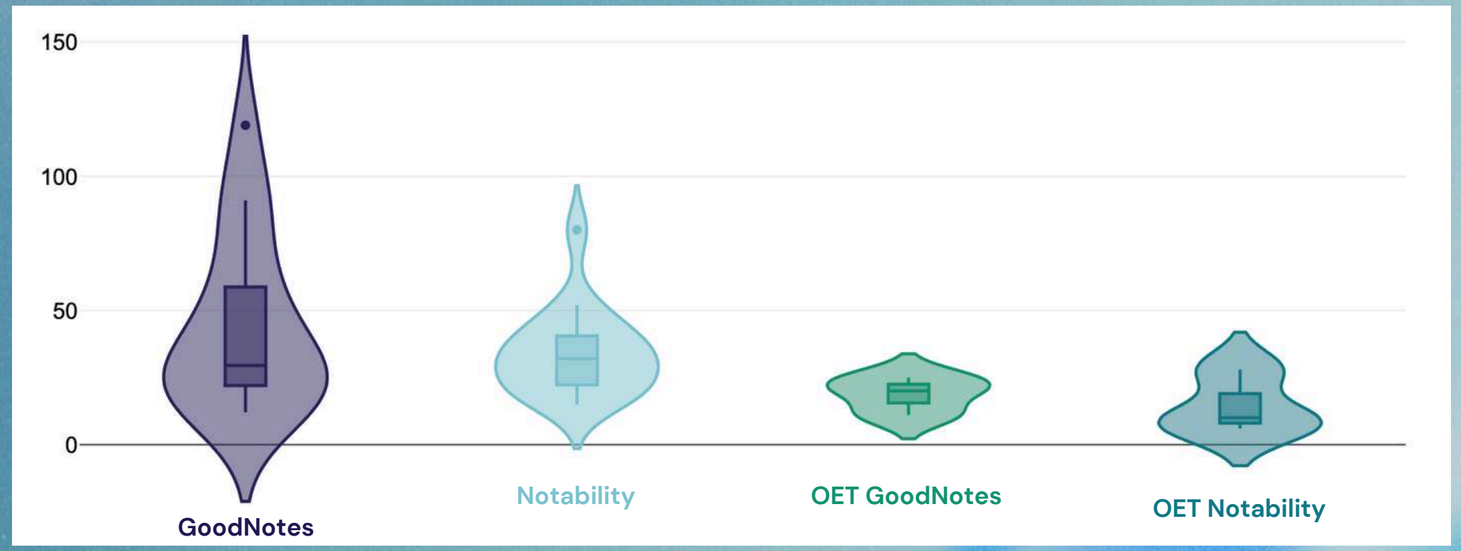
I risultati suggeriscono che gli utenti hanno impiegato un **tempo statisticamente significativamente maggiore rispetto all'OET** per eseguire il Task 1 su **Notability**.

Per GoodNotes invece **NON** è stata **riscontrata alcuna differenza statisticamente significativa** per quanto riguarda il tempo impiegato dagli utenti rispetto all'OET per eseguire il Task 1.

	Categoria	Media	Deviazione Standard	p-value
GoodNotes	Utenti	1:!5	0:43	0.156
	OET	0:38	0:20	
Notability	Utenti	1:21	0:35	0.0383
	OET	0:34	0:14	

Efficienza assoluta - Task 2

I risultati **NON** suggeriscono **alcuna differenza statisticamente significativa** rispetto all'OET nei tempi di esecuzione del Task 2 da parte degli utenti né su GoodNotes né su Notability.

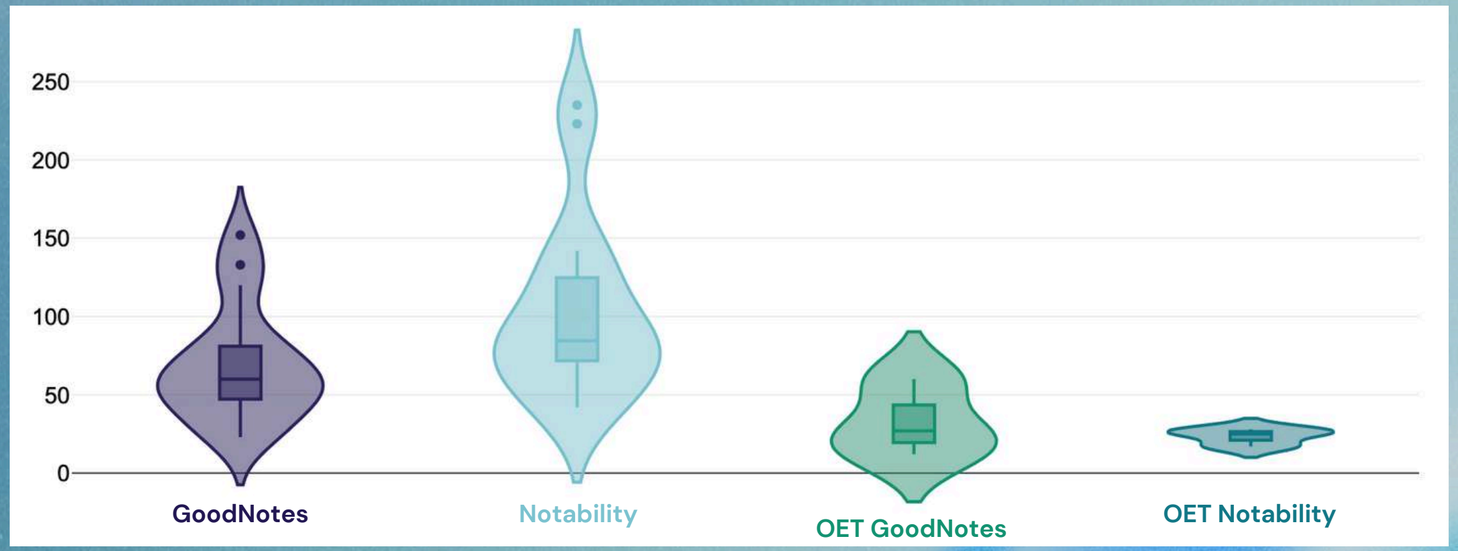


	Categoria	Media	Deviazione Standard	p-value
GoodNotes	Utenti	0:44	0:32	0.1909
	OET	0:19	0:07	
Notability	Utenti	0:34	0:15	0.0694
	OET	0:15	0:12	

Efficienza assoluta - Task 3

I risultati suggeriscono che gli utenti hanno impiegato un **tempo statisticamente significativamente maggiore rispetto all'OET** per eseguire il Task 3 su **Notability**.

Per GoodNotes invece **NON** è stata **riscontrata alcuna differenza statisticamente significativa** per quanto riguarda il tempo impiegato dagli utenti rispetto all'OET per eseguire il Task 3.



	Categoria	Media	Deviazione Standard	p-value
GoodNotes	Utenti	1:09	0:37	0.125
	OET	0:33	0:25	
Notability	Utenti	1:43	0:58	0.0306
	OET	0:23	0:06	

Espressioni facciali

Un'ulteriore misura della soddisfazione dell'utente è l'**analisi delle espressioni facciali**.
In questa fase sono state revisionate le riprese svolte durante lo svolgimento dei test utente per poter analizzare le espressioni facciali degli utenti **mentre interagivano con le due applicazioni**.

Questo ha permesso di raccogliere un loro **feedback aggiuntivo**, più "implicito" rispetto a quelli precedentemente forniti, e di individuare punti di criticità grazie alle **emozioni da loro espresse** suscitate dall'utilizzo dei sistemi.

In questa analisi vedremo:

CATEGORIZZAZIONE DELLE ESPRESSIONI

EMOZIONE PREVALENTE
(per ogni step di ciascuna task)

CONCLUSIONI
(calcolo dell'IRA*)

*Inter-Rater Agreement

Categorizzazione delle espressioni

POSITIVE

L'utente è **sereno** e non appare stressato.

NEUTRE

L'utente appare **perplesso** e/o particolarmente **concentrato** per compiere un compito (task).

NEGATIVE

L'utente è **visibilmente stressato**, **infastidito** e/o **confuso**.

Espressioni prevalenti - Task 1

	STEP 1 Crea una cartella	STEP 2 Crea un documento nella cartella	STEP 3 Disegna 3 fiori di colore diverso
GoodNotes	POSITIVE	NEUTRE	POSITIVE
Notability	NEUTRE	POSITIVE	POSITIVE

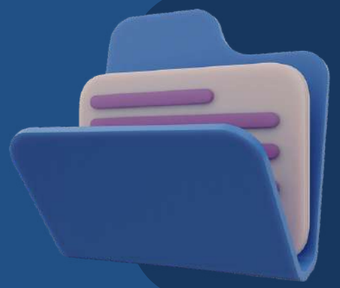
Espressioni prevalenti - Task 2

	STEP 1 Seleziona uno dei fiori	STEP 2 Crea una copia	STEP 3 Elimina l'originale
GoodNotes	POSITIVE	POSITIVE	POSITIVE
Notability	POSITIVE	POSITIVE	POSITIVE

Espressioni prevalenti - Task 3

	STEP 1 Apri un documento	STEP 2 Cerca la parola “prova”	STEP 3 Evidenzia e sottolinea
GoodNotes	POSITIVE	NEUTRE	POSITIVE
Notability	POSITIVE	NEGATIVE	POSITIVE

Considerazioni



TASK 1
Crea cartelle e documenti

Sono state registrate più **reazioni negative** per la **creazione di una cartella** su **Notability**. **GoodNotes** invece ha suscitato più **espressioni di confusione** per quanto riguarda la **creazione di un documento**.

TASK 2
Usa gli strumenti

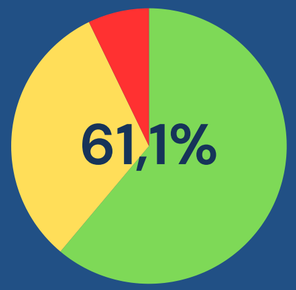
Non ci sono considerazioni particolari da fare; sono state registrate **reazioni positive** per entrambe le applicazioni.

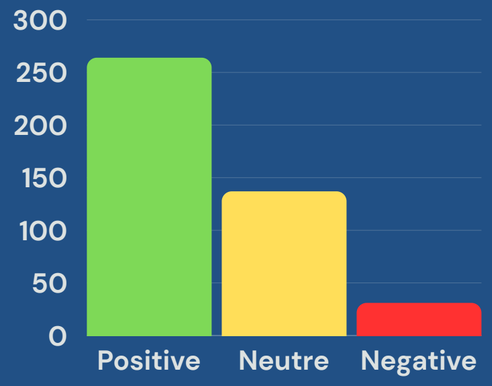
TASK 3
Cerca una parola nel documento

Entrambe le applicazioni, anche se Notability in maniera più grave, hanno suscitato **reazioni negative** negli utenti per la **ricerca di una parola all'interno di un documento**.

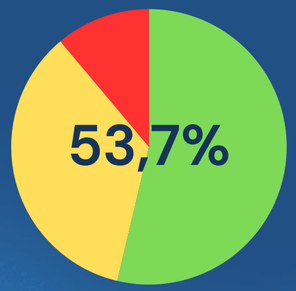
Riassunto Espressioni Registrate

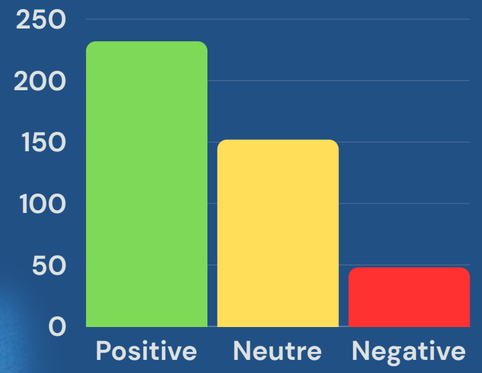
GoodNotes





Notability





ESPR. POSITIVE ESPR. NEUTRE ESPR. NEGATIVE

Dalle valutazioni raccolte risulta esserci una prevalenza di **espressioni positive** sia per GoodNotes che per Notability.

Inter-Rater Agreement (IRA)

Per calcolare il grado di accordo dei 3 valutatori che hanno svolto la valutazione delle espressioni facciali sono stati utilizzati come indicatori l'alpha di Krippendorf e il PO (Percentage Observer Agreement).
I dati raccolti a riguardo sono riportati di seguito:

ALPHA DI KRIPPENDORF*

calcolata attraverso la formula:

$$\alpha = \frac{D_0}{D_e}$$

con:
Do = disaccordo atteso
De = disaccordo dato dal caso atteso

α-Krippendorf = 0.504

PERCENTAGE OBSERVER AGREEMENT (PO)

Dopo aver raccolto tutte le opinioni dei valutatori, è stato calcolato il grado di accordo come

$$\frac{\text{numero di unità concordi}}{\text{numero totale di unità}}$$

PO = 0.59375

Il risultato ottenuto per il **PO non raggiunge la soglia minima di buon accordo** raccomandata del 67%, mentre l'**alpha di Krippendorf** calcolata rientra nella fascia **"accordo debole"** (0.4 < α ≤ 0.6).

*L'alpha di **Krippendorff** è stata calcolata utilizzando lo script chiamato ***krippendorf***

TEST UTENTE:

Considerazioni Finali

EFFICACIA

È stata rilevata una differenza statisticamente significativa nell'esecuzione del Task 3: gli utenti hanno commesso un numero significativamente maggiore di errori con Notability rispetto che con GoodNotes.

EFFICIENZA

È stata rilevata una differenza statisticamente significativa nell'esecuzione del Task 3: gli utenti hanno impiegato un tempo significativamente maggiore con Notability rispetto che con GoodNotes.

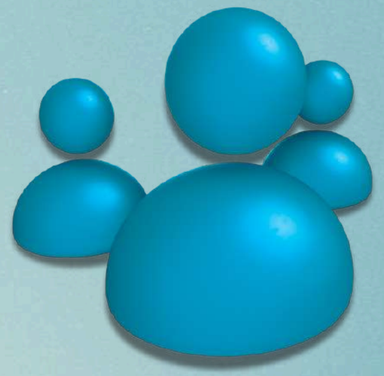
EFFICACIA ASSOLUTA

È stata rilevata una differenza statisticamente significativa nell'esecuzione del Task 3 nell'applicazione di Notability rispetto alla soglia ottimale di errore.

EFFICIENZA ASSOLUTA

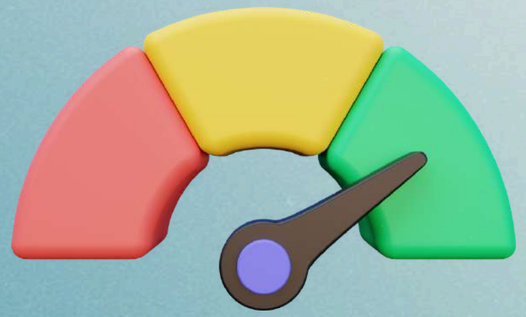
Sono state rilevate differenze statisticamente significative nell'esecuzione dei Task 1 e 3 sull'applicazione di Notability rispetto agli OET.

QUESTIONARIO



Utenti Coinvolti

UEQ



NPS

Conclusioni

Overview

Per misurare la soddisfazione degli utenti è stato somministrato un questionario per ciascuna delle due applicazioni. Gli utenti coinvolti sono stati 32, tra cui i 3 esperti di dominio coinvolti nella Valutazione Euristica e i 13 utenti coinvolti nel Test Utente.

Il questionario ha la seguente struttura:

- una prima sezione per la raccolta di **dati di carattere generale** (sesso, età, occupazione, ecc)
- una seconda sezione per il **UEQ** (User Experience Questionnaire), composto da 26 domande
- una domanda finale per il calcolo del **NPS** (Net Promoter Score)

UEQ

26 domande alle quali va assegnato un valore da 1 a 7*.

Scelto in quanto:

- coinvolge diverse aree semantiche
- sono disponibili diversi strumenti di calcolo e una traduzione italiana online ufficiali
- è breve e facile da compilare

NPS

1 domanda alla quale bisogna assegnare un punteggio da 0 a 10.

Scelto in quanto:

- è semplice
- permette confronti nel tempo
- tiene conto dell'Acquiescence Bias**

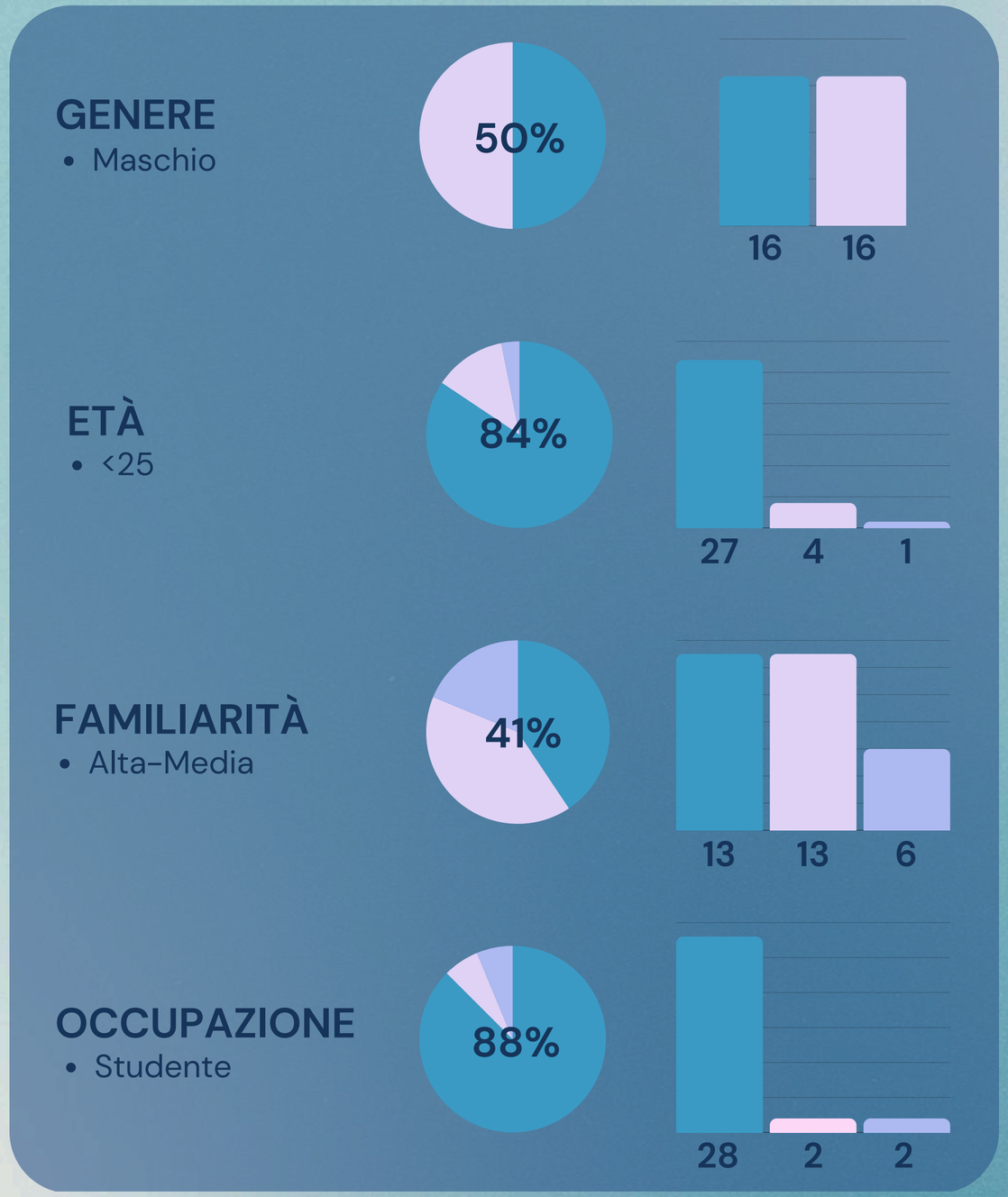
*siccome questo tipo di scala non minimizza il Central Tendency Bias, (cioè la tendenza a selezionare l'opzione centrale), abbiamo utilizzato una scala a 8 valori, contando i 2 valori centrali come se fossero un unico valore, in modo da renderla compatibile con quella originale ma potendo comunque capire la tendenza positiva e negativa
Acquiescence Bias = tendenza a preferire un'opzione di risposta positiva

Utenti coinvolti

ID UTENTE	GENERE	FASCIA D'ETÀ	OCCUPAZIONE	FAMILIARITÀ
ED1	M	<25	Studente	Alta
ED2	M	<25	Studente	Media
ED3	F	<25	Studente	Media
U4	F	<25	Studente	Bassa
U5	F	<25	Studente	Alta
U6	M	>50	Lavoratore	Bassa
U7	M	<25	Studente	Alta
U8	F	tra 25 e 50	Lavoratore	Bassa
U9	M	<25	Studente	Media
U10	M	<25	Studente	Media
U11	M	<25	Studente	Alta
U12	F	<25	Studente	Media
U13	M	tra 25 e 50	Studente	Alta
U14	F	<25	Studente	Media
U15	F	<25	Studente-Lavoratore	Bassa
U16	F	<25	Studente	Bassa

ID UTENTE	GENERE	FASCIA D'ETÀ	OCCUPAZIONE	FAMILIARITÀ
Q17	F	<25	Studente	Media
Q18	F	<25	Studente	Alta
Q19	F	<25	Studente	Alta
Q20	F	<25	Studente	Media
Q21	M	<25	Studente	Media
Q22	M	<25	Studente-Lavoratore	Alta
Q23	M	<25	Studente	Alta
Q24	M	Tra 25 e 50	Studente	Media
Q25	M	<25	Studente	Alta
Q26	M	<25	Studente	Media
Q27	F	<25	Studente	Media
Q28	F	<25	Studente	Alta
Q29	F	<25	Studente	Media
Q30	M	Tra 25 e 50	Studente	Bassa
Q31	F	<25	Studente	Alta
Q32	M	<25	Studente	Alta

Composizione del campione



I valutatori coinvolti sono:

- equamente distribuiti per quanto riguarda il genere
- prevalentemente di età **inferiore ai 25 anni**
- perlopiù con **familiarità medio-alta**
- per la maggior parte **studenti**.

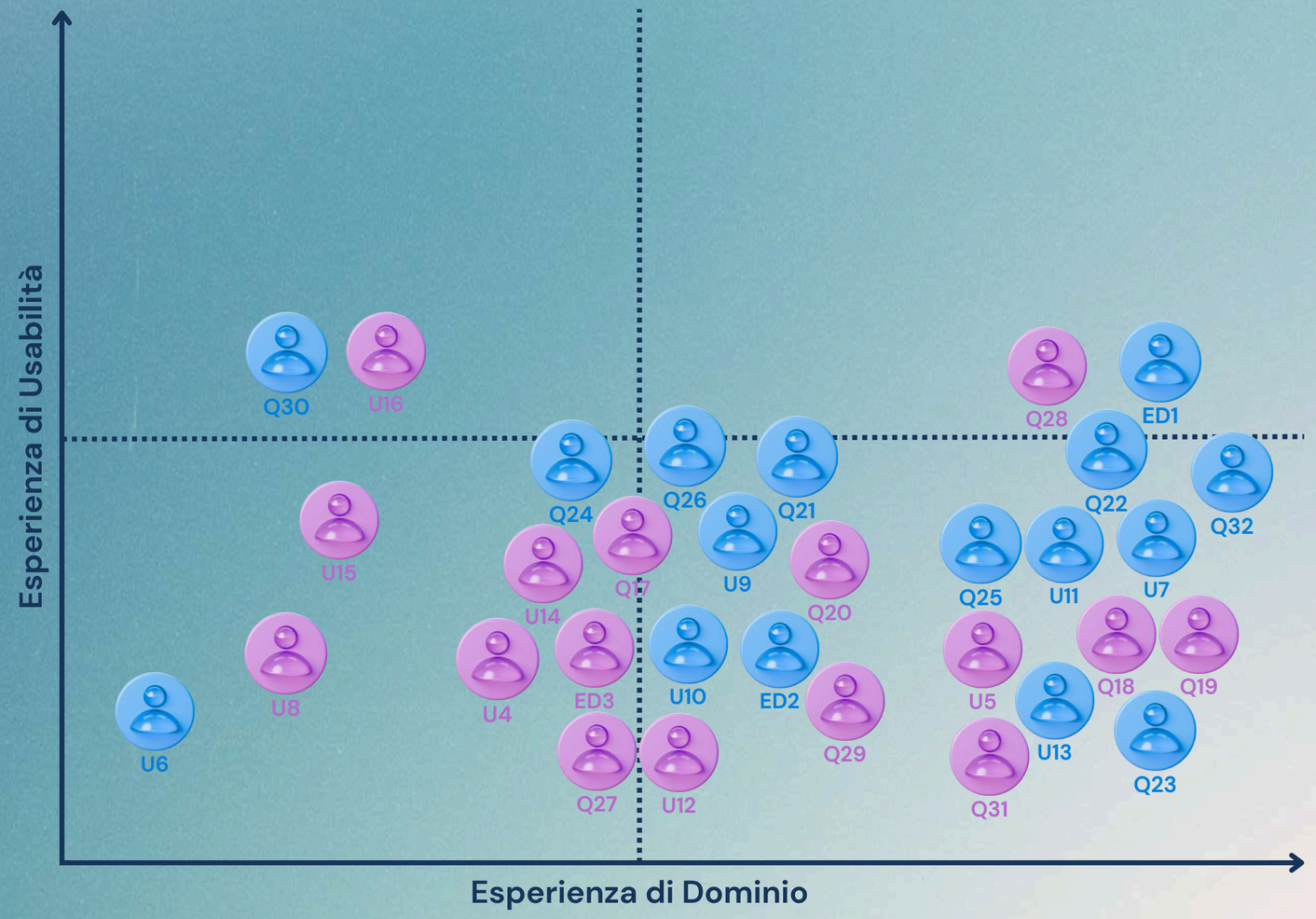
Matrice di Expertise

La matrice di expertise permette di vedere

- l'**esperienza di usabilità**
- l'**esperienza di dominio**

di ciascun utente coinvolto.

In particolare nel nostro caso si può osservare una conoscenza dell'**usabilità medio-bassa** e una conoscenza **equamente distribuita** del **dominio**.



UEQ - User Experience Questionnaire

Lo UEQ è un questionario di **26 domande**, categorizzate in **6 scale** (riportate nella slide seguente).
Per ogni domanda vengono proposte **due caratteristiche in contrapposizione** tra loro a cui viene richiesto di assegnare un valore da 1 a 7*, dove 1 indica massimo accordo con l'opzione di sinistra e 7 con quello di destra.

I risultati di ciascuna domanda sono rappresentati graficamente nelle prossime slides con dei boxplot, che permettono di fare un confronto diretto delle due applicazioni.

Successivamente verranno analizzate in dettaglio alcune domande (quelle che abbiamo ritenuto più significative) utilizzando un test di Wilcoxon con un livello di confidenza del 95%, cioè le differenze sono ritenute statisticamente significative con un p-value inferiore al 5%.

*siccome questo tipo di scala non minimizza il Central Tendency Bias, (cioè la tendenza a selezionare l'opzione centrale), abbiamo utilizzato una scala a 8 valori, contando i 2 valori centrali come se fossero un unico valore, in modo da renderla compatibile con quella originale ma potendo comunque capire la tendenza positiva e negativa

Scala UEQ

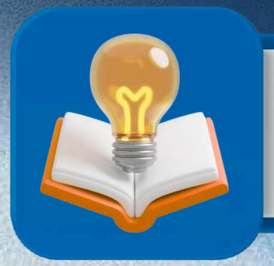
Attrattività
Impressione generale del prodotto

Apprendibilità
Facilità di confidenza al prodotto

Efficienza
Sforzo per raggiungere obiettivi

Controllabilità
Controllo, sicurezza, prevedibilità

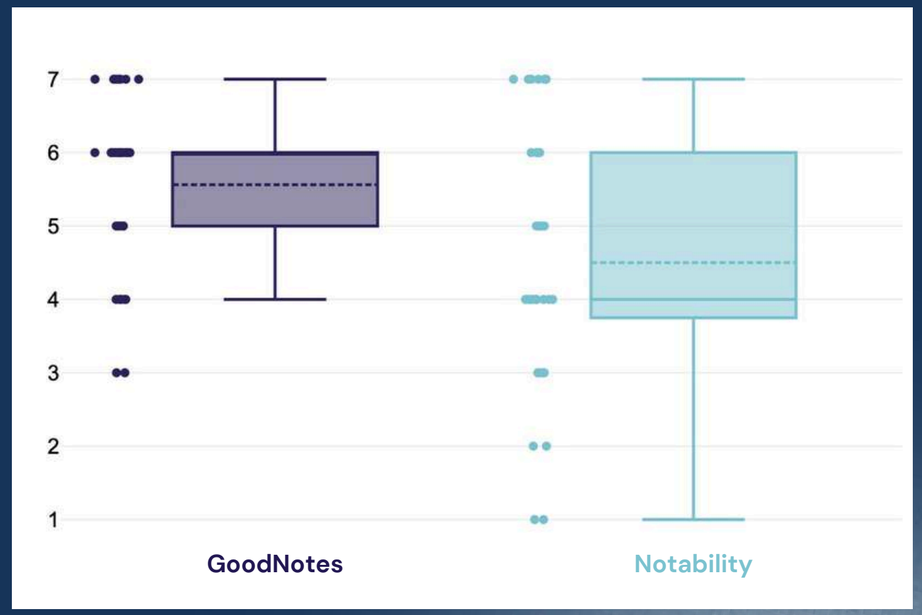
Stimolazione
Motivazione all'utilizzo



Originalità
Design distinguibile dagli altri

DOMANDA 1

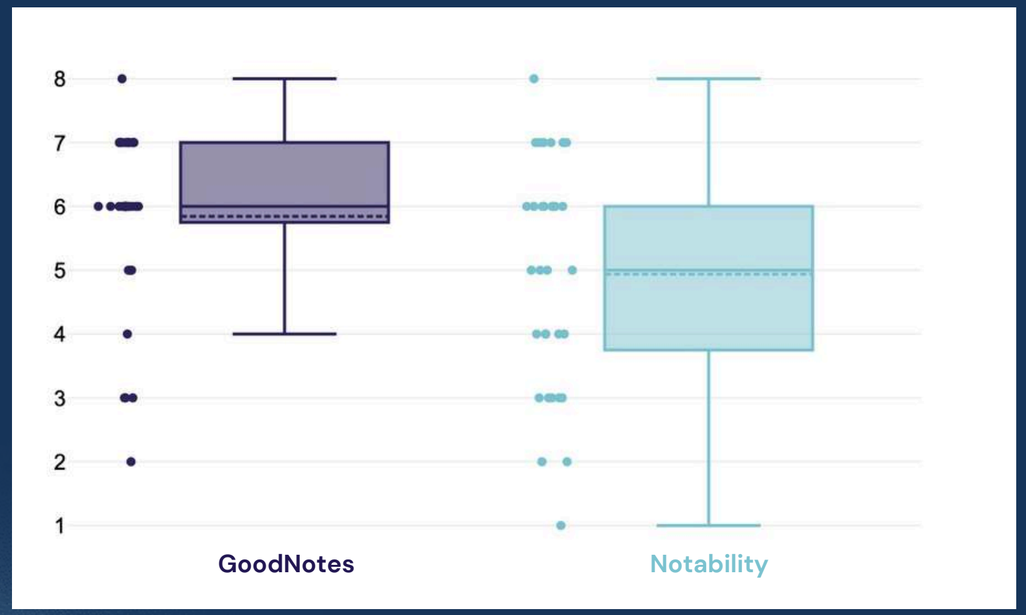
Fastidioso (1) – Piacevole (8)



APPLICAZIONE	Baffo Inferiore	Primo Quartile	Mediana	Terzo Quartile	Baffo Superiore	Media
GoodNotes	4	5	6	6	7	5,56
Notability	1	3,75	4	6	7	4,5

DOMANDA 2

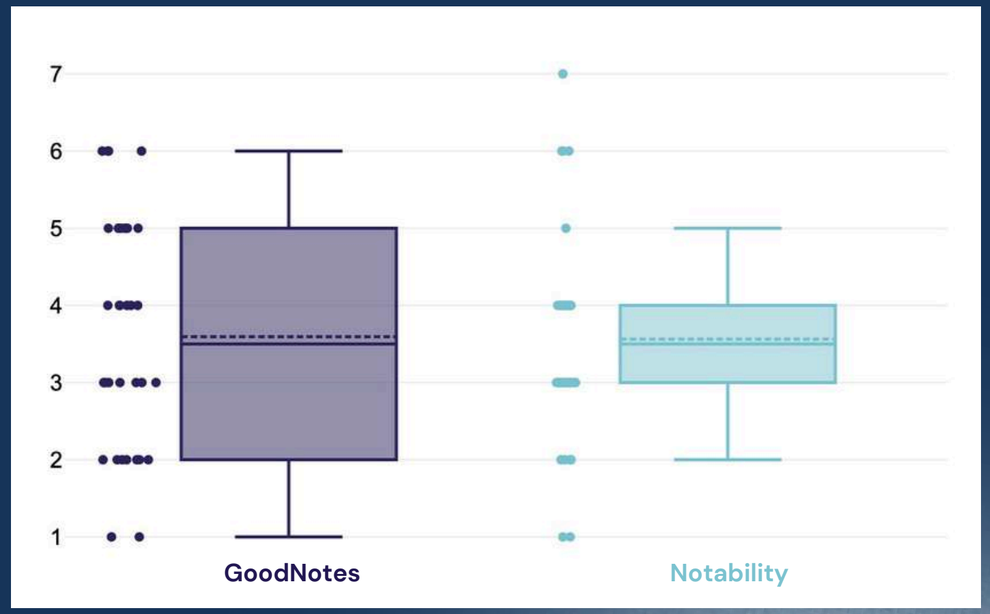
Incomprensibile (1) – Comprensibile (8)



APPLICAZIONE	Baffo Inferiore	Primo Quartile	Mediana	Terzo Quartile	Baffo Superiore	Media
GoodNotes	4	5,75	6	7	8	5,84
Notability	1	3,75	5	6	8	4,94

DOMANDA 3

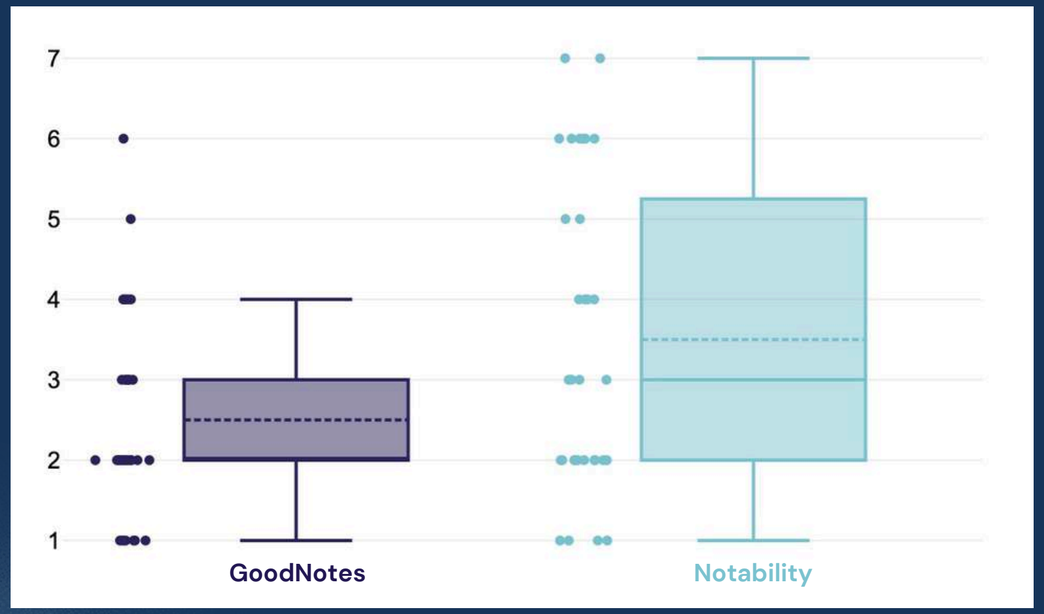
Creativo (1) – Privo di fantasia (8)



APPLICAZIONE	Baffo Inferiore	Primo Quartile	Mediana	Terzo Quartile	Baffo Superiore	Media
GoodNotes	1	2	3,5	5	6	3,59
Notability	2	3	3,5	4	5	3,56

DOMANDA 4

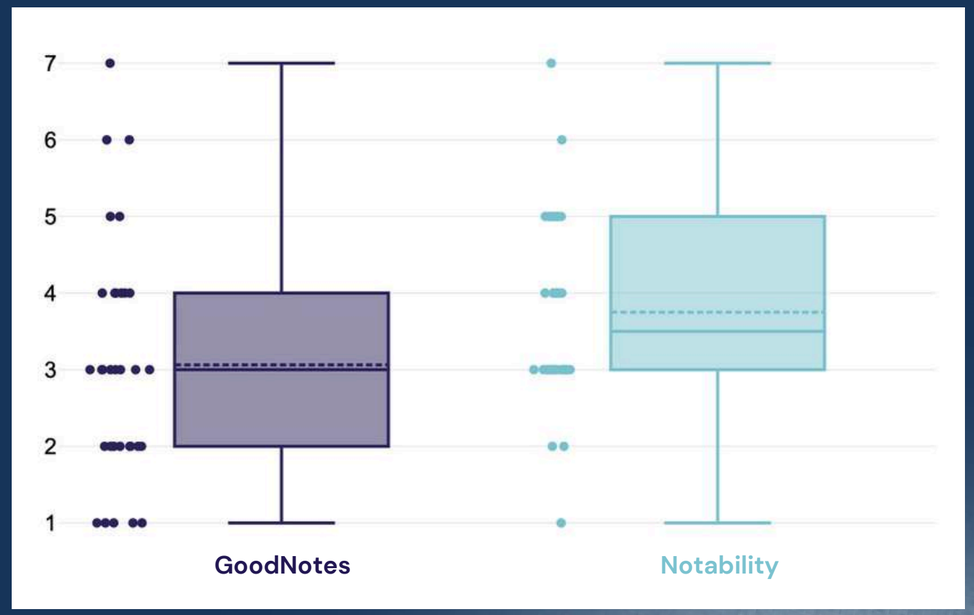
Facile da apprendere (1) – Difficile da apprendere (8)



APPLICAZIONE	Baffo Inferiore	Primo Quartile	Mediana	Terzo Quartile	Baffo Superiore	Media
GoodNotes	1	2	2	3	4	2,5
Notability	1	2	3	5,25	7	3,5

DOMANDA 5

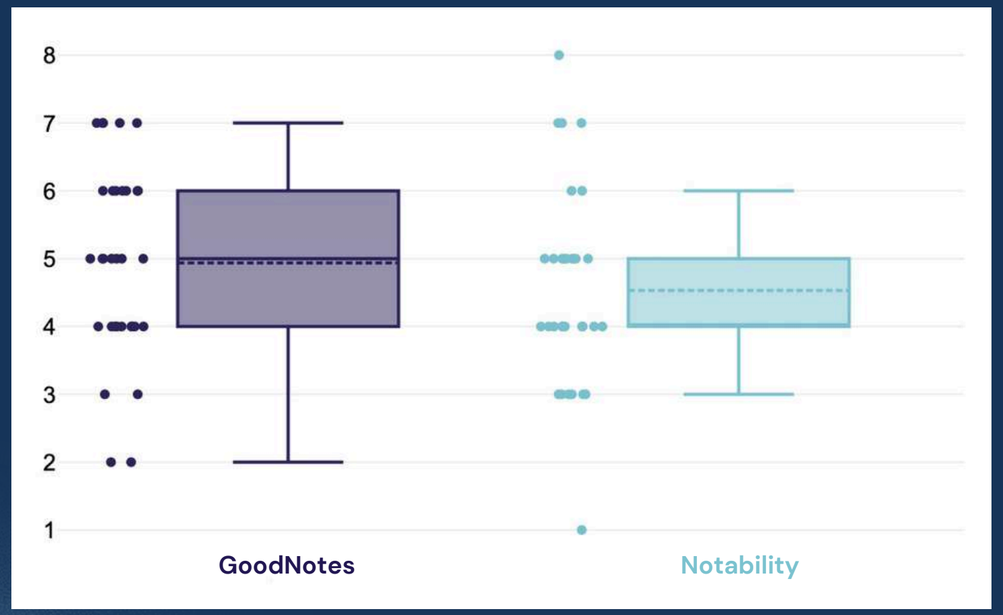
Di grande valore (1) - Di poco valore (8)



APPLICAZIONE	Baffo Inferiore	Primo Quartile	Mediana	Terzo Quartile	Baffo Superiore	Media
GoodNotes	1	2	3	4	7	3,06
Notability	1	3	3,5	5	7	3,75

DOMANDA 6

Noioso (1) - Appassionante (8)



APPLICAZIONE	Baffo Inferiore	Primo Quartile	Mediana	Terzo Quartile	Baffo Superiore	Media
GoodNotes	2	4	5	6	7	4,94
Notability	3	4	4	5	6	4,53

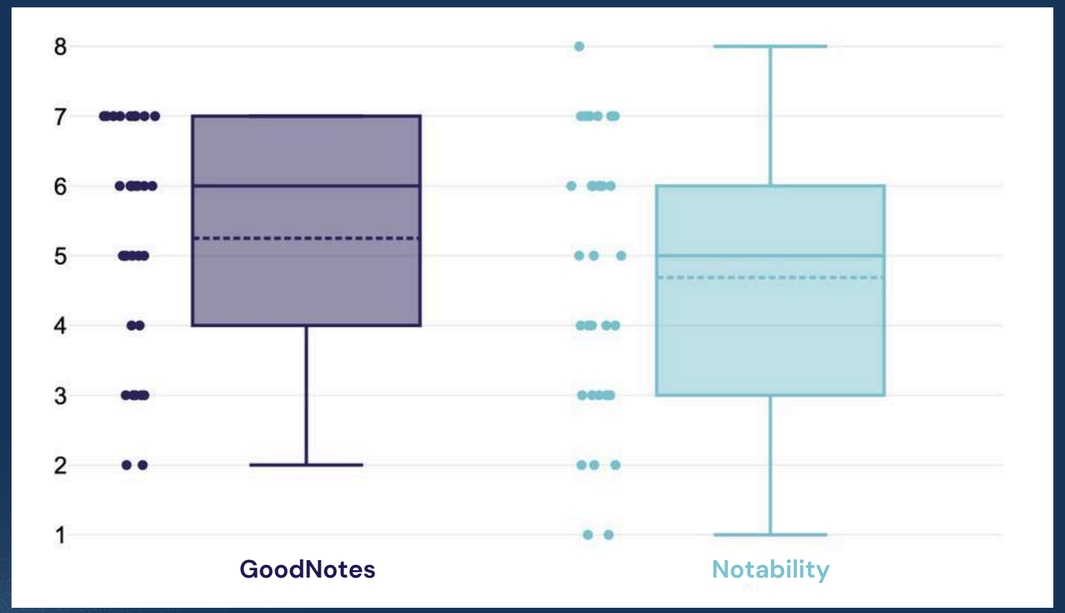
DOMANDA 7

Non interessante (1) – Interessante (8)

APPLICAZIONE	Baffo Inferiore	Primo Quartile	Mediana	Terzo Quartile	Baffo Superiore	Media
GoodNotes	3	4	5,5	7	8	5,44
Notability	1	4	5	6	7	4,72

DOMANDA 8

Imprevedibile (1) – Prevedibile (8)



APPLICAZIONE	Baffo Inferiore	Primo Quartile	Mediana	Terzo Quartile	Baffo Superiore	Media
GoodNotes	2	4	6	7	7	5,25
Notability	1	3	5	6	8	4,69

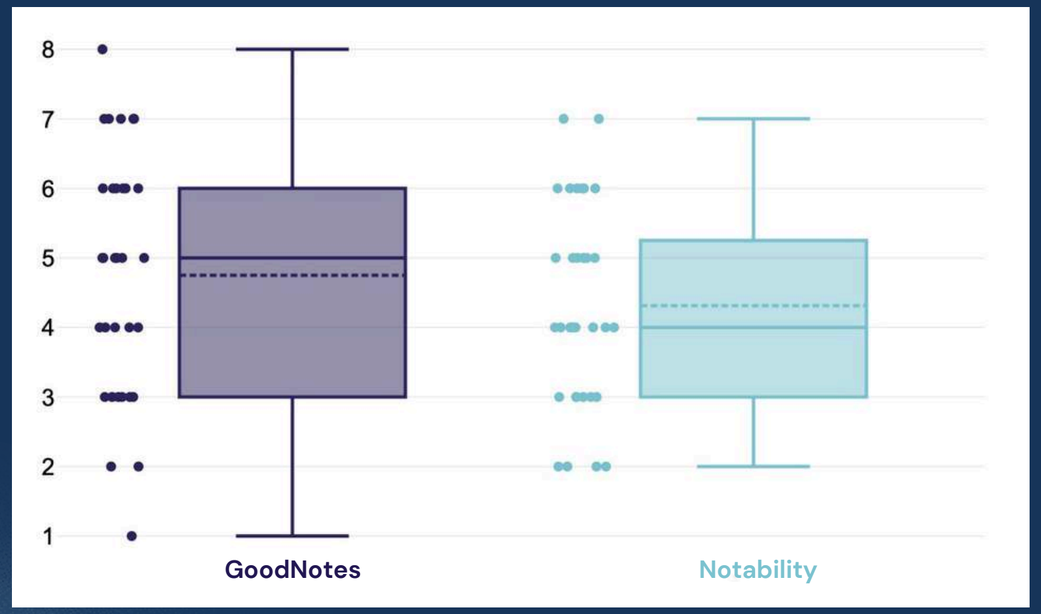
DOMANDA 9

Veloce (1) - Lento (8)

APPLICAZIONE	Baffo Inferiore	Primo Quartile	Mediana	Terzo Quartile	Baffo Superiore	Media
GoodNotes	1	2	3	4	7	3,09
Notability	1	2	3	3,25	5	3

DOMANDA 10

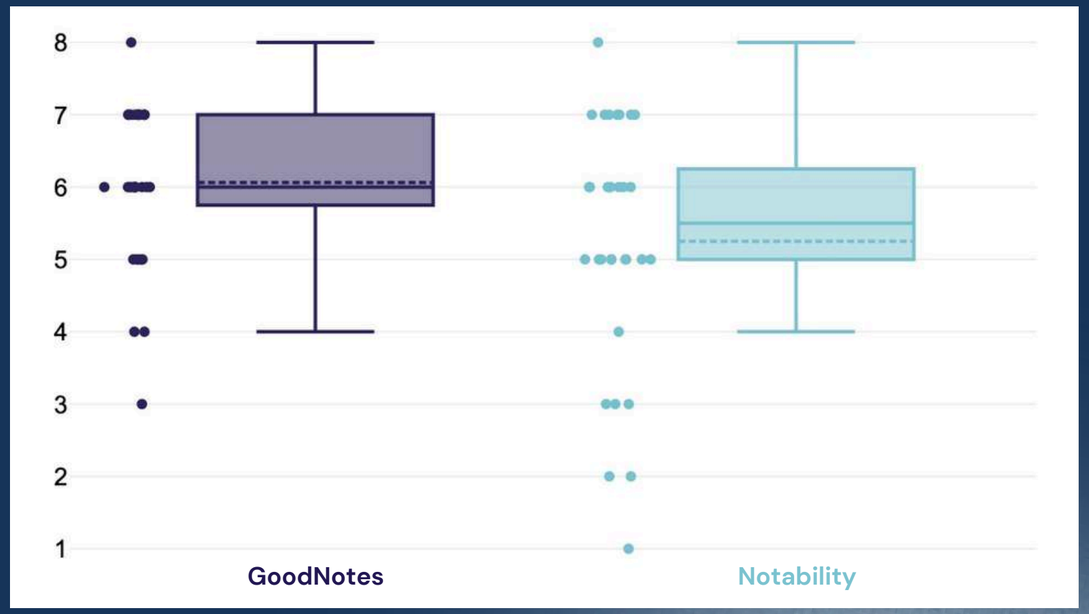
Originale (1) - Convenzionale (8)



APPLICAZIONE	Baffo Inferiore	Primo Quartile	Mediana	Terzo Quartile	Baffo Superiore	Media
GoodNotes	1	3	5	6	8	4,75
Notability	2	3	4	5,25	7	4,31

DOMANDA 11

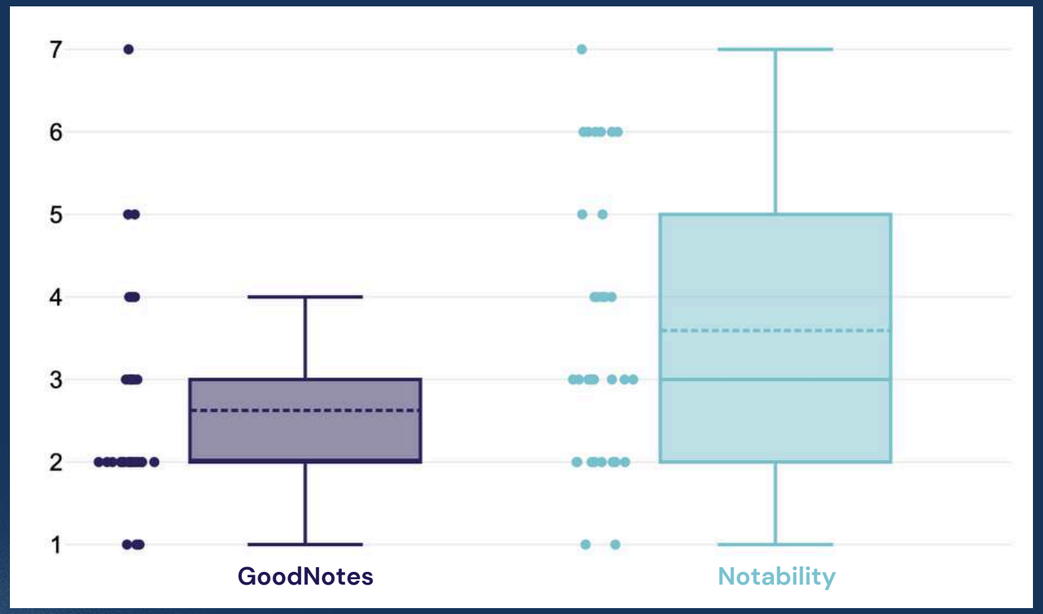
Ostruttivo (1) - Di supporto (8)



APPLICAZIONE	Baffo Inferiore	Primo Quartile	Mediana	Terzo Quartile	Baffo Superiore	Media
GoodNotes	4	5,75	6	7	8	6,06
Notability	4	5	5,5	6,25	8	5,25

DOMANDA 12

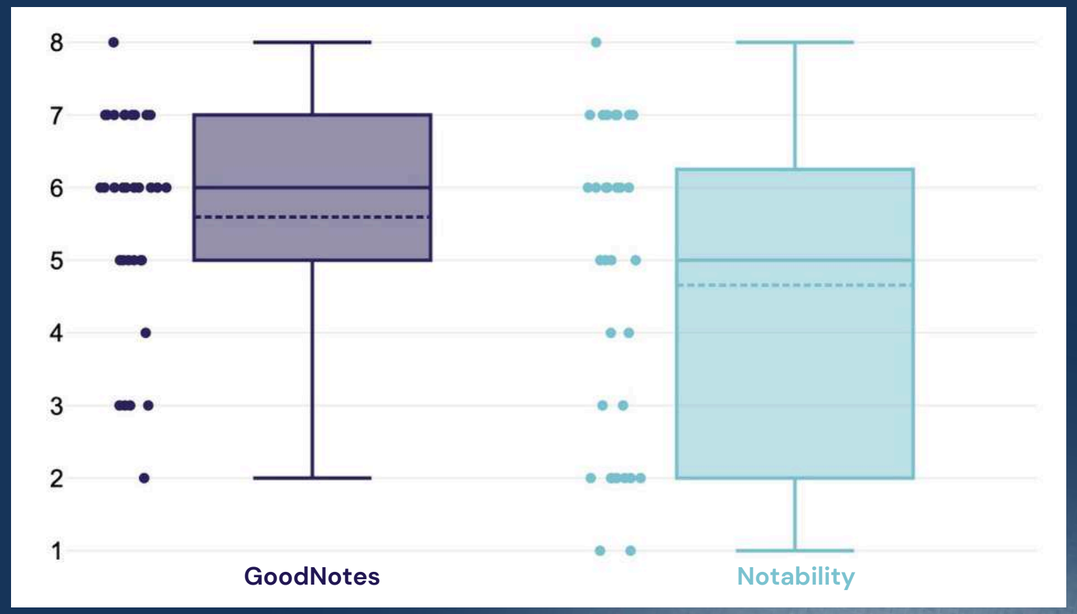
Buono (1) - Scarso (8)



APPLICAZIONE	Baffo Inferiore	Primo Quartile	Mediana	Terzo Quartile	Baffo Superiore	Media
GoodNotes	1	2	2	3	4	2,63
Notability	1	2	3	5	7	3,59

DOMANDA 13

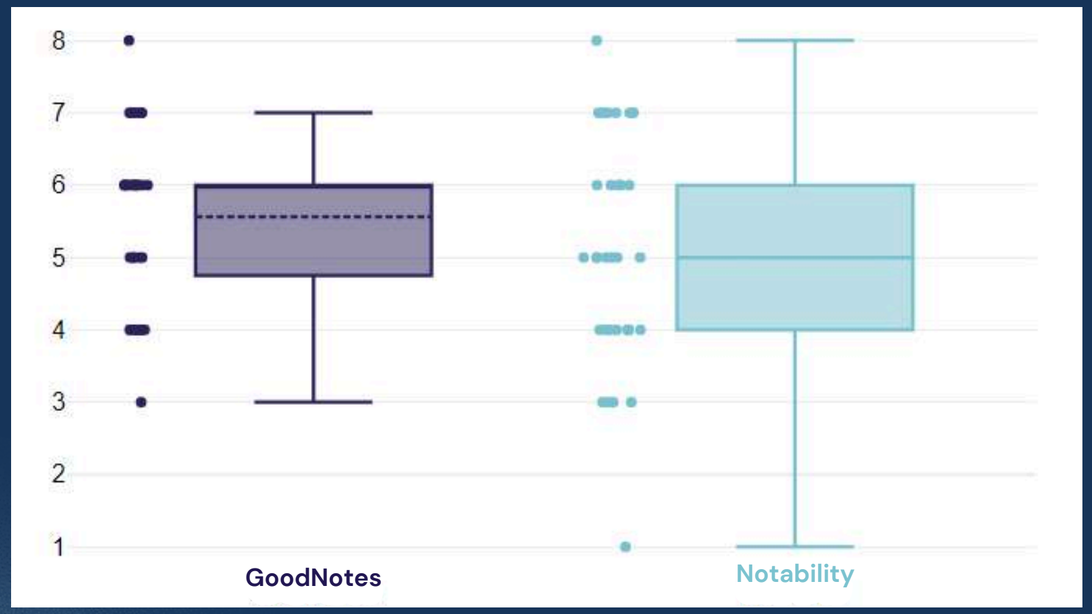
Complicato (1) - Facile (8)



APPLICAZIONE	Baffo Inferiore	Primo Quartile	Mediana	Terzo Quartile	Baffo Superiore	Media
GoodNotes	2	5	6	7	8	5,59
Notability	1	2	5	6,25	8	4,63

DOMANDA 14

Repellente (1) - Attraente(8)



APPLICAZIONE	Baffo Inferiore	Primo Quartile	Mediana	Terzo Quartile	Baffo Superiore	Media
GoodNotes	3	4,75	6	6	7	5,56
Notability	1	4	5	6	8	5

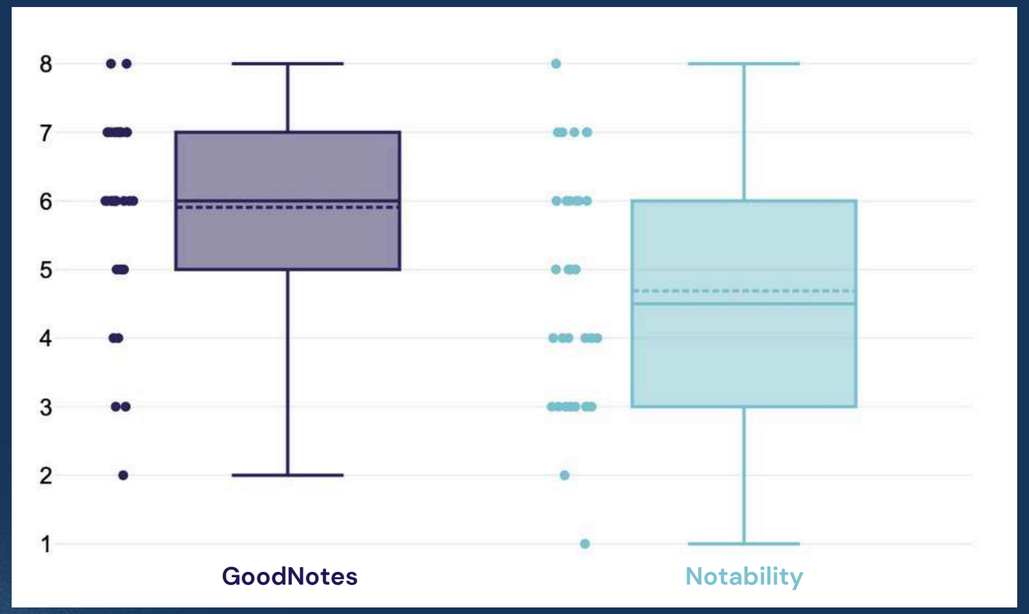
DOMANDA 15

Usuale (1) - Moderno (8)

APPLICAZIONE	Baffo Inferiore	Primo Quartile	Mediana	Terzo Quartile	Baffo Superiore	Media
GoodNotes	1	4	5	6	8	4,97
Notability	2	4	5	6	8	5,03

DOMANDA 16

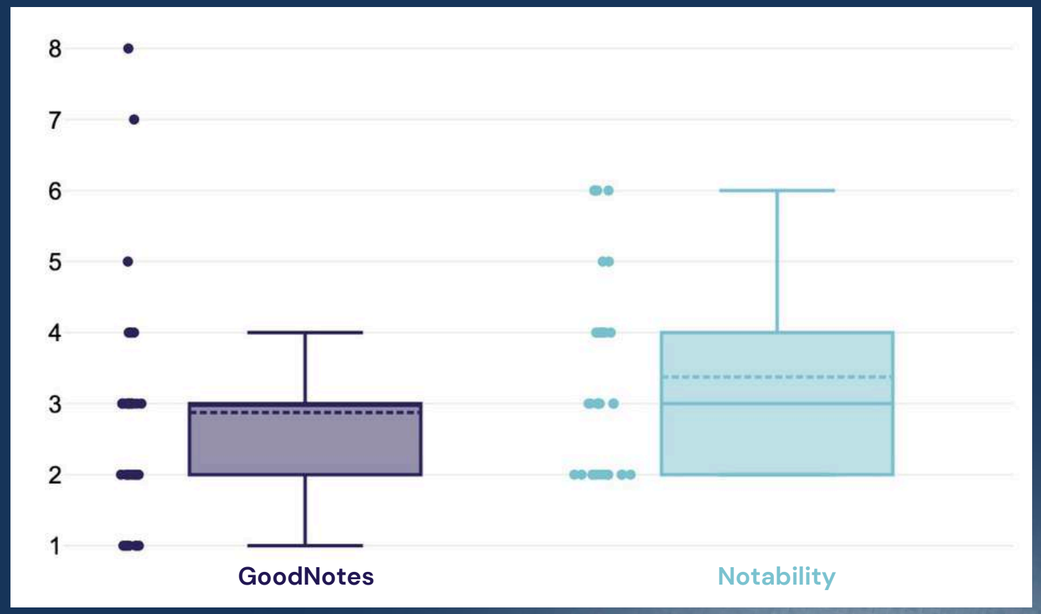
Sgradevole (1) - Gradevole (8)



APPLICAZIONE	Baffo Inferiore	Primo Quartile	Mediana	Terzo Quartile	Baffo Superiore	Media
GoodNotes	2	5	6	7	8	5,91
Notability	1	3	4,5	6	8	4,69

DOMANDA 17

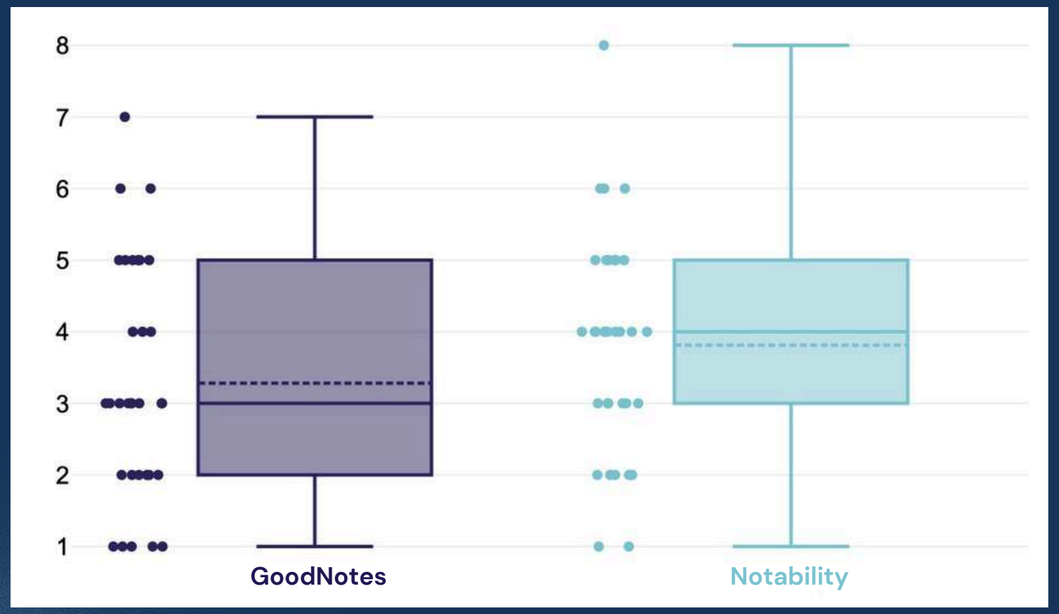
Sicuro (1) - Insicuro (8)



APPLICAZIONE	Baffo Inferiore	Primo Quartile	Mediana	Terzo Quartile	Baffo Superiore	Media
GoodNotes	1	2	3	3	4	2,84
Notability	2	2	3	4	6	3,47

DOMANDA 18

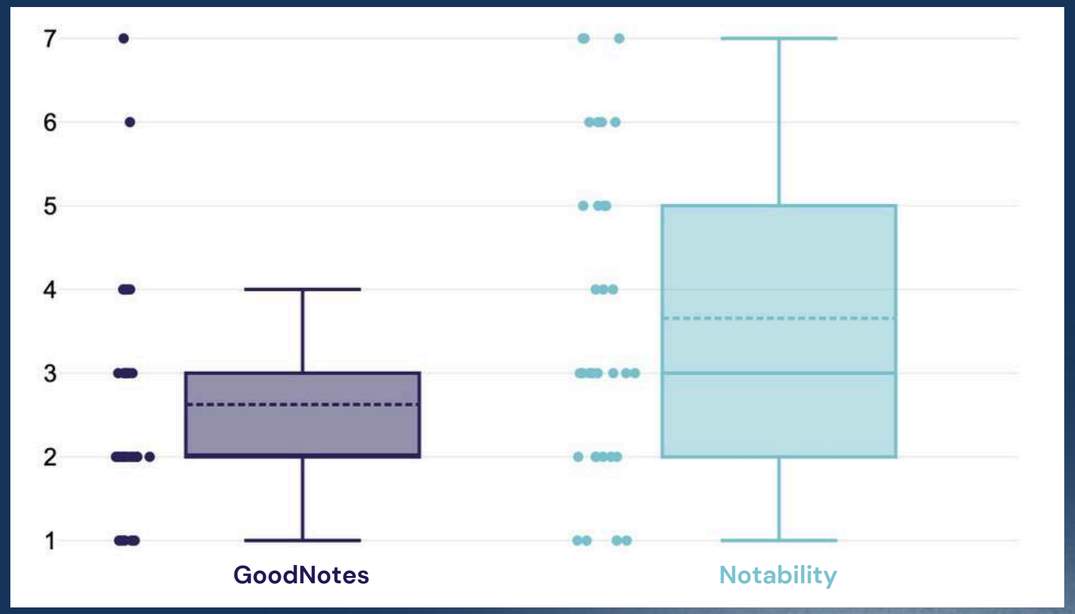
Attivante (1) - Soporifero (8)



APPLICAZIONE	Baffo Inferiore	Primo Quartile	Mediana	Terzo Quartile	Baffo Superiore	Media
GoodNotes	1	2	3	5	7	3,28
Notability	1	3	4	5	8	3,81

DOMANDA 19

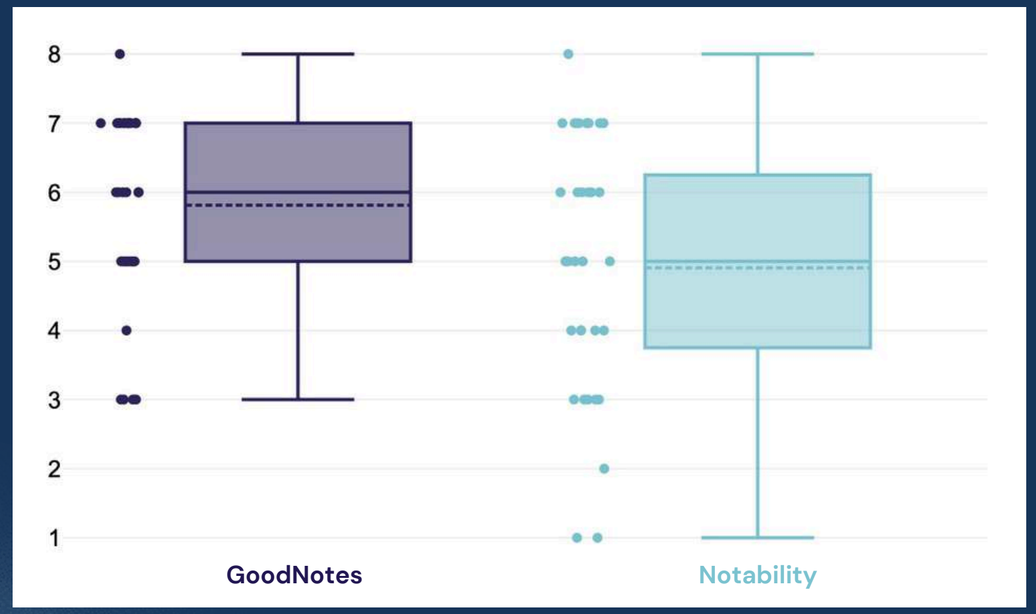
Conforme alle aspettative (1) - Non conforme alle aspettative (8)



APPLICAZIONE	Baffo Inferiore	Primo Quartile	Mediana	Terzo Quartile	Baffo Superiore	Media
GoodNotes	1	2	2	3	4	2,63
Notability	1	2	3	5	7	3,69

DOMANDA 20

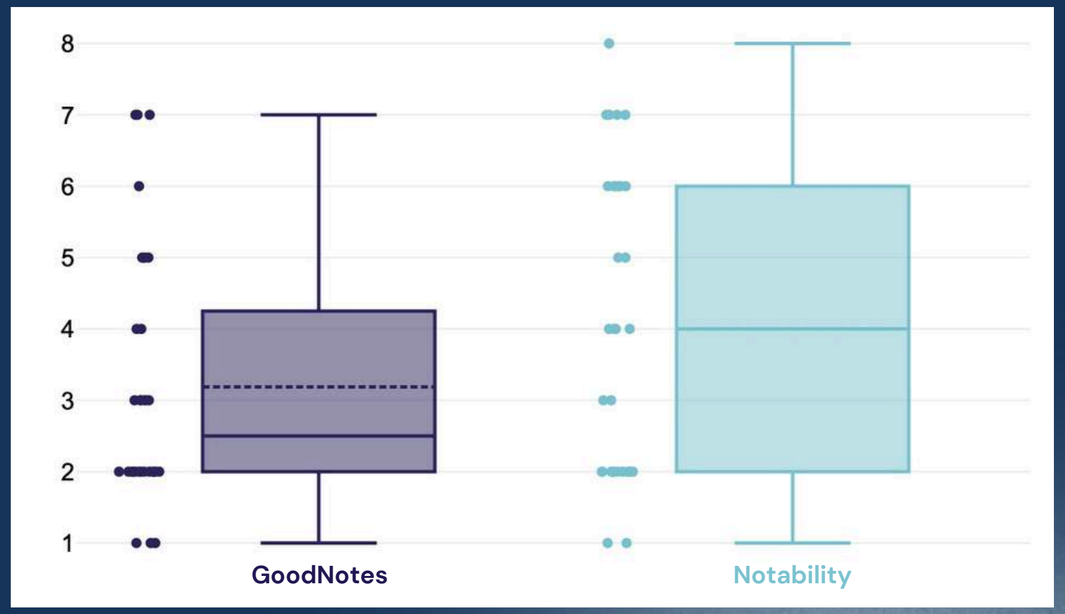
Inefficiente (1) - Efficiente (8)



APPLICAZIONE	Baffo Inferiore	Primo Quartile	Mediana	Terzo Quartile	Baffo Superiore	Media
GoodNotes	3	5	6	7	8	5,81
Notability	1	3,75	5	6,25	8	4,91

DOMANDA 21

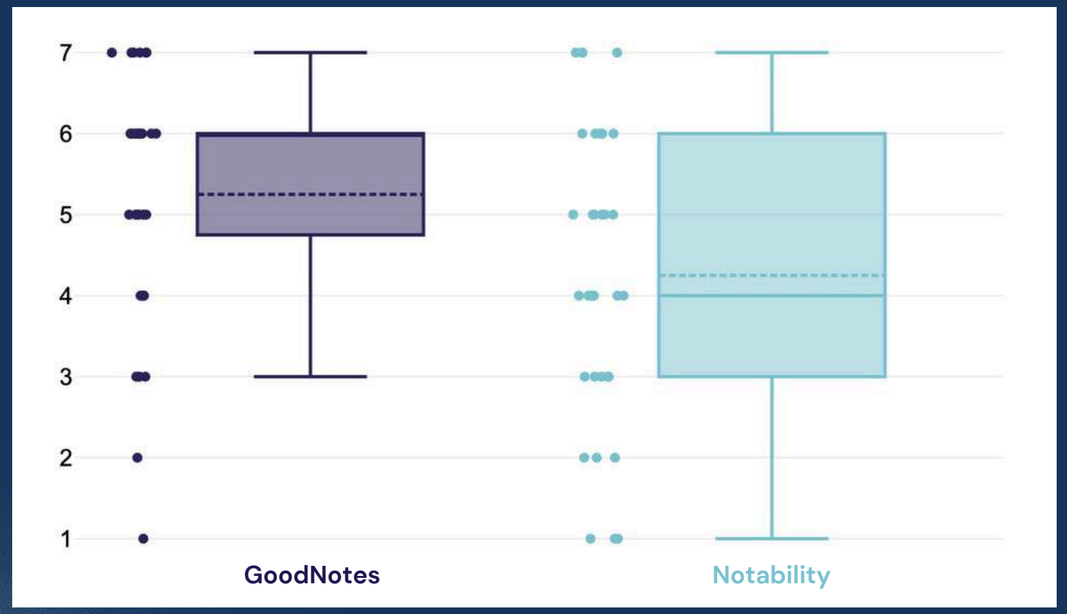
Chiaro (1) - Confuso (8)



APPLICAZIONE	Baffo Inferiore	Primo Quartile	Mediana	Terzo Quartile	Baffo Superiore	Media
GoodNotes	1	2	2,5	4,25	7	3,19
Notability	1	2	4	6	8	4

DOMANDA 22

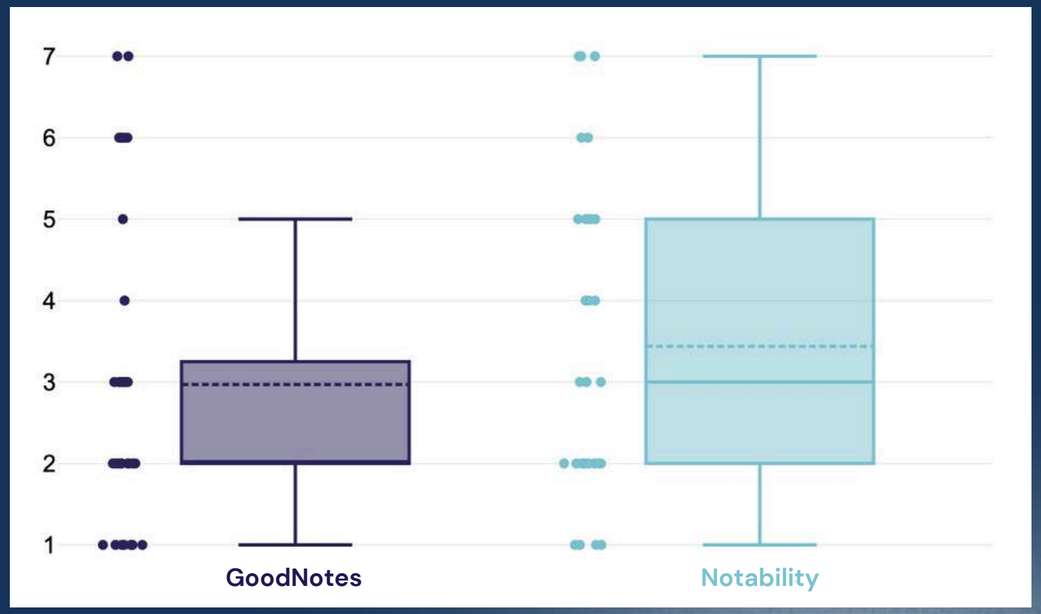
Non pragmatico (1) - Pragmatico (8)



APPLICAZIONE	Baffo Inferiore	Primo Quartile	Mediana	Terzo Quartile	Baffo Superiore	Media
GoodNotes	3	4,75	6	6	7	5,25
Notability	1	3	4	6	7	4,25

DOMANDA 23

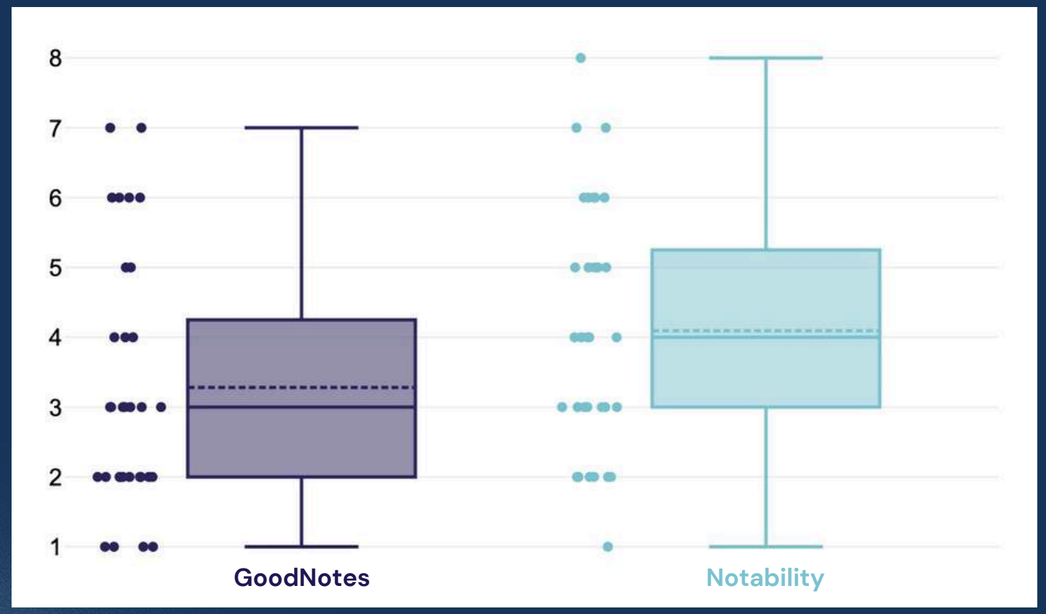
Ordinato (1) - Sovraccarico (8)



APPLICAZIONE	Baffo Inferiore	Primo Quartile	Mediana	Terzo Quartile	Baffo Superiore	Media
GoodNotes	1	2	2	3,25	5	2,97
Notability	1	2	3	5	7	3,44

DOMANDA 24

Invitante (1) - Non invitante (8)



APPLICAZIONE	Baffo Inferiore	Primo Quartile	Mediana	Terzo Quartile	Baffo Superiore	Media
GoodNotes	1	2	3	4,25	7	3,28
Notability	1	3	4	5,25	8	4,09

DOMANDA 25

Congeniale (1) - Ostile (8)

APPLICAZIONE	Baffo Inferiore	Primo Quartile	Mediana	Terzo Quartile	Baffo Superiore	Media
GoodNotes	1	2	3	4	5	3,19
Notability	1	3	4	5	7	3,97

DOMANDA 26

Conservativo (1) - Innovativo (8)

APPLICAZIONE	Baffo Inferiore	Primo Quartile	Mediana	Terzo Quartile	Baffo Superiore	Media
GoodNotes	1	4	5	6	7	4,81
Notability	1	4	5	6	7	4,78

Confronto tra servizi

Nelle prossime slides analizzeremo più nel dettaglio 4 domande del questionario, ovvero:

- DOMANDA 2: Incomprensibile - Comprensibile
- DOMANDA 8: Prevedibile - Imprevedibile
- DOMANDA 16: Sgradevole - Piacevole
- DOMANDA 20: Inefficiente - Efficiente

Sulle risposte a tali domande applicheremo il **test dei ranghi con segno di Wilcoxon per campioni accoppiati**.
Utilizzeremo un test a **doppia coda** con **intervallo di confidenza del 95%**, ovvero consideriamo le differenze statisticamente significative quando otteniamo un p-value inferiore al 5%.

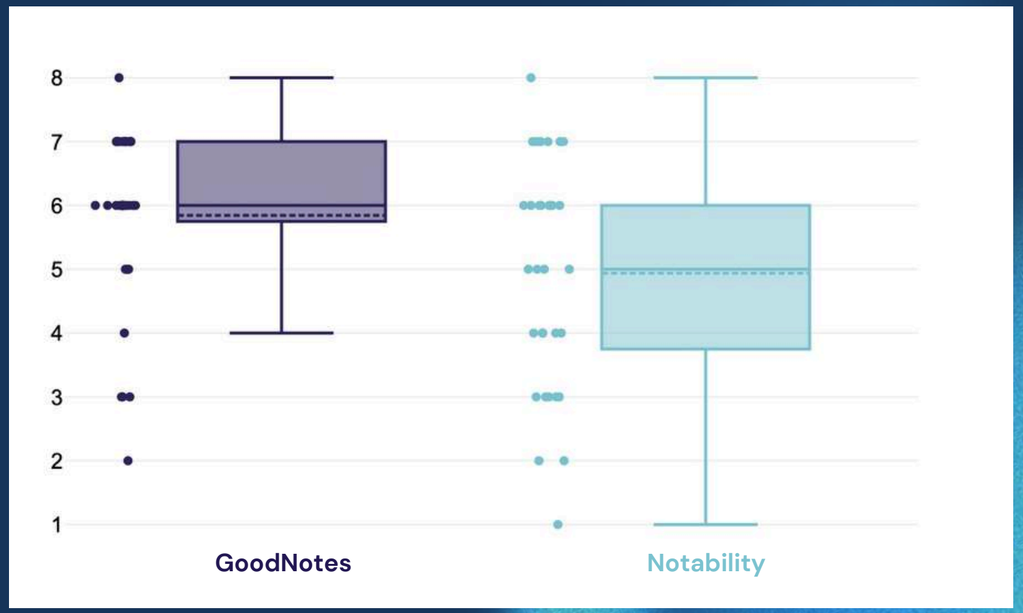
Comprensibilità

Il Test di Wilcoxon è stato eseguito sui 32 partecipanti al fine di determinare se esistono differenze statisticamente significative nella percezione della **comprensibilità** tra i due servizi esaminati.

Dai risultati sono stati ottenuti:

- z = -1.799
- p-value = 0.072

NON è quindi **stata trovata alcuna differenza statisticamente significativa** per quanto riguarda la comprensibilità delle due applicazioni.



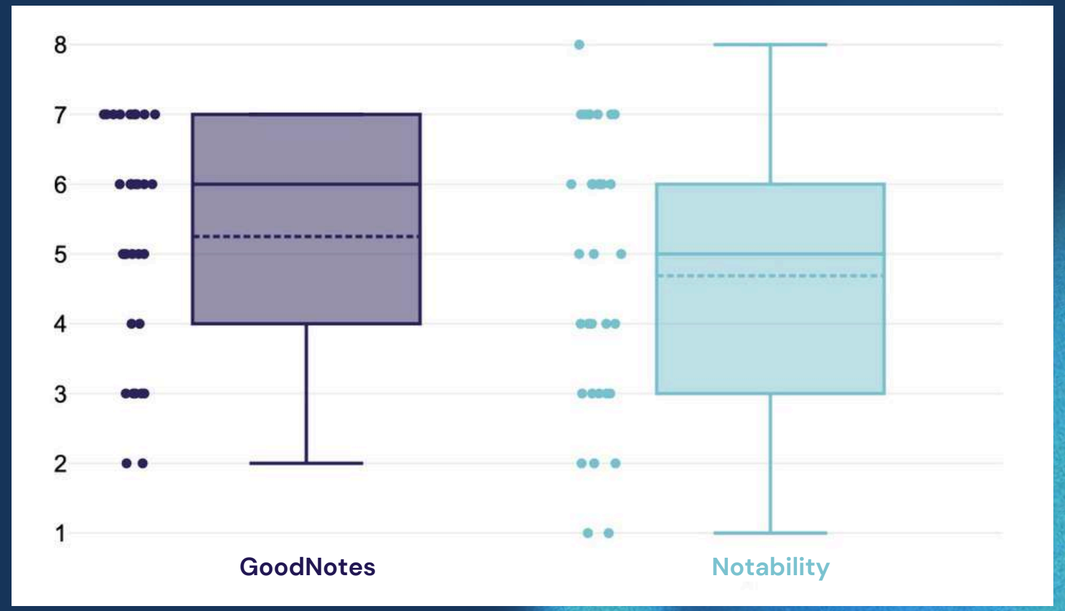
Predevibilità

Il Test di Wilcoxon è stato eseguito sui 32 partecipanti al fine di determinare se esistono differenze statisticamente significative nella percezione della **prevedibilità** tra i due servizi esaminati.

Dai risultati sono stati ottenuti:

- z = −1.1
- p-value = 0.271

NON è quindi **stata trovata alcuna differenza statisticamente significativa** per quanto riguarda la prevedibilità delle due applicazioni.



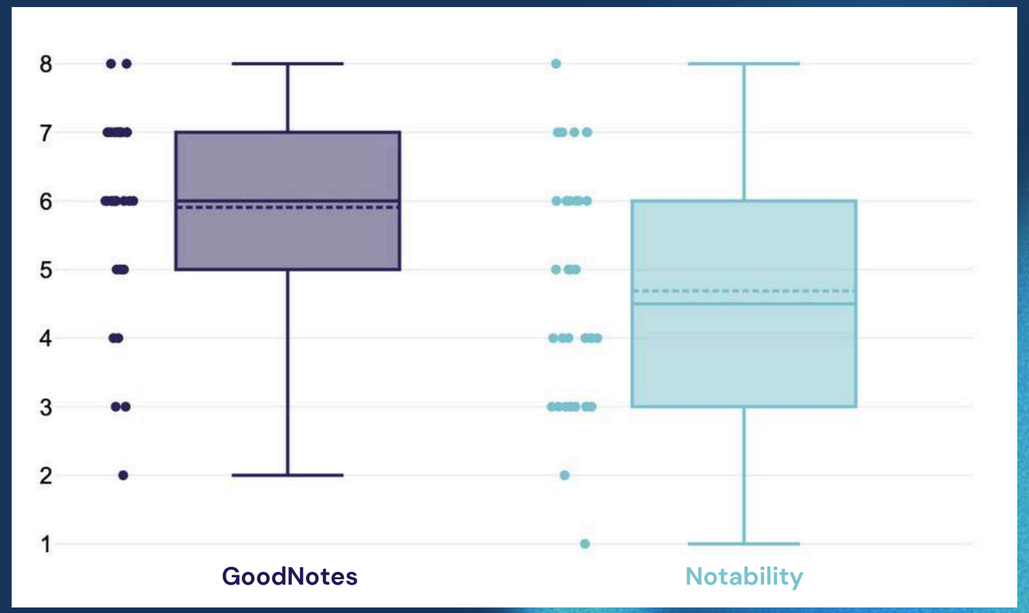
Piacevolezza

Il Test di Wilcoxon è stato eseguito sui 32 partecipanti al fine di determinare se esistono differenze statisticamente significative nella percezione della **piacevolezza** tra i due servizi esaminati.

Dai risultati sono stati ottenuti:

- z = −2.415
- p-value = 0.016

È stata trovata una differenza statisticamente significativa per quanto riguarda la piacevolezza delle due applicazioni: **GoodNotes** viene percepita **più piacevole** rispetto a Notability.



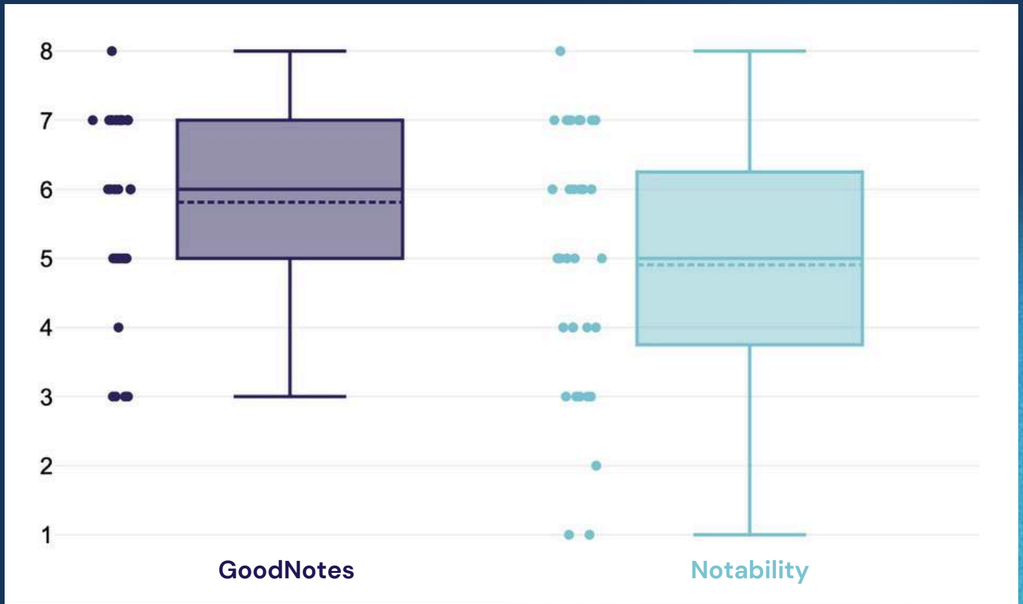
Efficienza

Il Test di Wilcoxon è stato eseguito sui 32 partecipanti al fine di determinare se esistono differenze statisticamente significative nella percezione dell'**efficienza** tra i due servizi esaminati.

Dai risultati sono stati ottenuti:

- z = -2.108
- p-value = 0.035

È stata trovata una differenza statisticamente significativa per quanto riguarda l'efficienza delle due applicazioni: **GoodNotes** è ritenuta **più efficiente** rispetto a Notability.



Analisi UEQ

Nelle slides successive riportiamo le considerazioni fatte dopo aver analizzato i dati tramite il tool ufficiale* del UEQ , che classifica le domande nelle 6 scale precedentemente citate nella seguente maniera:

ATTRATTIVITÀ

D1: Fastidioso-Piacevole
D12: Buono-Scarso
D14: Repellente-Attraente
D16: Sgradevole-Gradevole
D21: Invitante-Non invitante
D25: Congeniale-Ostile

APPRENDIBILITÀ

D2: Incomprensibile-Comprensibile
D4: Facile da apprendere-Difficile da apprendere
D13: Complicato-Facile
D21: Chiaro-Confuso

EFFICIENZA

D9: Veloce-Lento
D20: Inefficiente-Efficiente
D22: Non pragmatico-Pragmatico
D23: Ordinato-Sovraccarico

CONTROLLABILITÀ

D8: Imprevedibile-Prevedibile
D11: Ostruttivo-Di supporto
D17: Sicuro-Insicuro
D19: Conforme alle aspettative-Non conforme alle aspettative

STIMOLAZIONE

D5: Di grande valore-Di poco valore
D6: Noioso-Appassionante
D7: Non interessante-Interessante
D18: Attivante-Soporifero

ORIGINALITÀ

D3: Creativo-Privo di fantasia
D10: Originale-Convenzionale
D15: Usuale-Moderno
D26: Conservativo-Innovativo

*https://www.ueq-online.org/

Analisi delle domande - GoodNotes

Numero	Media	Varianza	Dev. Standard	Valore Sinistro	Valore Destro	Scala
1	1.6	1.4	1.2	increscioso	gradito	Attrattività
2	1.8	2.1	1.4	incomprensibile	comprensibile	Apprendibilità
3	0.4	2.2	1.5	creativo	privo di fantasia	Originalità
4	1.5	1.5	1.2	facile da capire	difficile da capire	Apprendibilità
5	0.9	2.4	1.6	di grande valore	di poco valore	Stimolazione
6	0.9	2.0	1.4	noioso	appassionante	Stimolazione
7	1.4	2.1	1.5	non interessante	interessante	Stimolazione
8	1.3	2.6	1.6	imprevedibile	prevedibile	Controllabilità
9	0.9	3.0	1.7	veloce	lento	Efficienza
10	-0.8	3.1	1.8	originale	convenzionale	Originalità
11	2.1	1.2	1.1	ostruttiva	di supporto	Controllabilità
12	1.4	1.7	1.3	bene	male	Attrattività
13	1.6	2.2	1.5	complicato	facile	Apprendibilità

Numero	Media	Varianza	Dev. Standard	Valore Sinistro	Valore Destro	Scala
14	1.6	1.5	1.2	repellente	attraente	Attrattività
15	1.0	2.9	1.7	usuale	moderno	Originalità
16	1.9	2.1	1.4	sgradevole	piacevole	Attrattività
17	1.2	2.2	1.5	sicuro	insicuro	Controllabilità
18	0.7	2.7	1.7	attivante	soporifero	Stimolazione
19	1.4	1.9	1.4	aspettativo	non aspettativo	Controllabilità
20	1.8	2.0	1.4	inefficiente	efficiente	Efficienza
21	0.8	3.2	1.8	chiaro	confuso	Apprendibilità
22	1.3	2.4	1.5	non pragmatico	pragmatico	Efficienza
23	1.0	3.6	1.9	ordinato	sovraccarico	Efficienza
24	0.7	3.3	1.8	attrattivo	non attrattivo	Attrattività
25	0.8	2.4	1.6	simpatico	antipatico	Attrattività
26	0.8	2.6	1.6	conservativo	innovativo	Originalità

Analisi delle domande - Notability

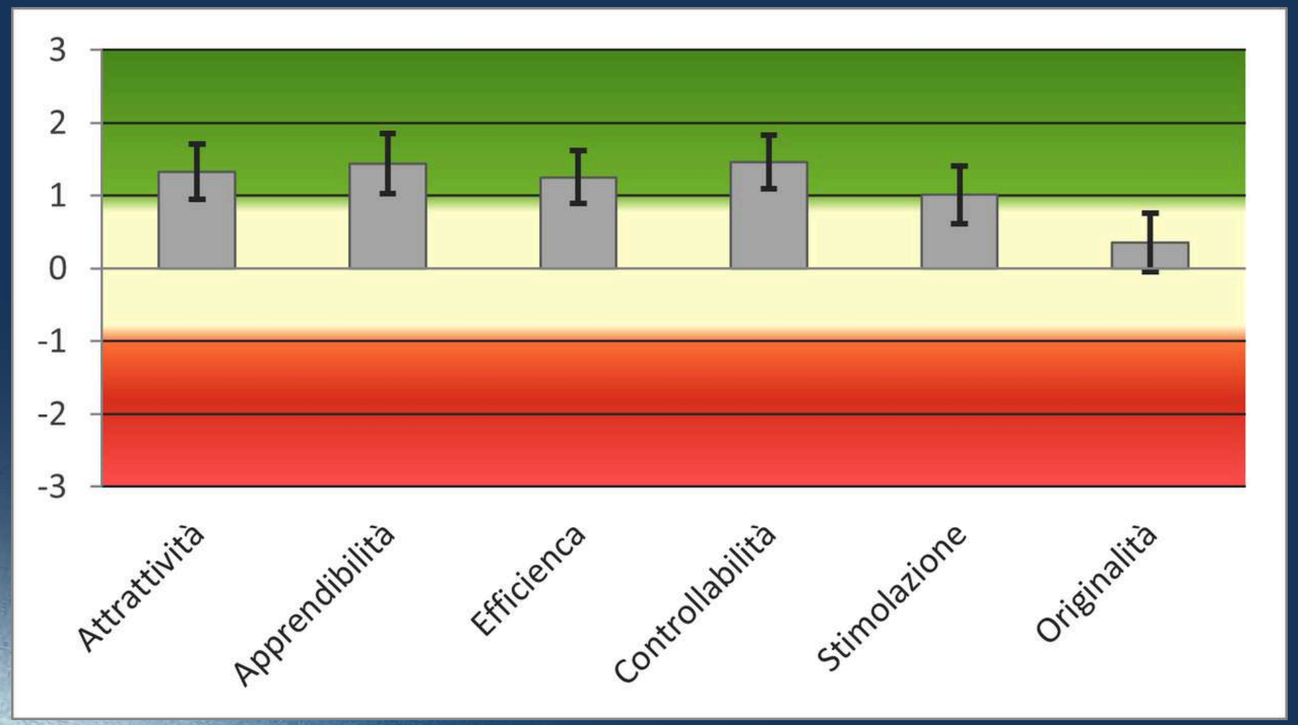
Numero	Media	Varianza	Dev. Standard	Valore Sinistro	Valore Destro	Scala
1	0.5	3.1	1.8	increscioso	gradito	Attrattività
2	0.9	3.2	1.8	incomprensibile	comprensibile	Apprendibilità
3	0.4	1.9	1.4	creativo	privo di fantasia	Originalità
4	0.5	3.7	1.9	facile da capire	difficile da capire	Apprendibilità
5	0.3	1.5	1.2	di grande valore	di poco valore	Stimolazione
6	0.5	2.1	1.5	noioso	appassionante	Stimolazione
7	0.7	2.1	1.4	non interessante	interessante	Stimolazione
8	0.7	3.9	2.0	imprevedibile	prevedibile	Controllabilità
9	1.0	1.5	1.2	veloce	lento	Efficienca
10	-0.3	2.2	1.5	originale	convenzionale	Originalità
11	1.3	2.9	1.7	ostruttiva	di supporto	Controllabilità
12	0.4	2.8	1.7	bene	male	Attrattività
13	0.6	4.9	2.2	complicato	facile	Apprendibilità

Numero	Media	Varianza	Dev. Standard	Valore Sinistro	Valore Destro	Scala
14	1.0	2.5	1.6	repellente	attraente	Attrattività
15	1.0	2.2	1.5	usuale	moderno	Originalità
16	0.7	3.1	1.7	sgradevole	piacevole	Attrattività
17	0.5	1.9	1.4	sicuro	insicuro	Controllabilità
18	0.2	2.5	1.6	attivante	soporifero	Stimolazione
19	0.3	3.8	1.9	aspettativo	non aspettativo	Controllabilità
20	0.9	3.5	1.9	inefficiente	efficiente	Efficienca
21	0.0	4.6	2.1	chiaro	confuso	Apprendibilità
22	0.3	3.4	1.8	non pragmatico	pragmatico	Efficienca
23	0.6	3.7	1.9	ordinato	sovraccarico	Efficienca
24	-0.1	3.2	1.8	attrattivo	non attrattivo	Attrattività
25	0.0	2.7	1.7	simpatico	antipatico	Attrattività
26	0.8	1.9	1.4	conservativo	innovativo	Originalità

Analisi dei sottogruppi - GoodNotes

Sottogruppo	Media	Dev. Standard	No.	Confidenza	I.C. (p = 0.05)	
Attrattività	1.323 ↑	1.102	32	0.382	0.941	1.705
Apprendibilità	1.438 ↑	1.200	32	0.416	1.022	1.853
Efficienca	1.250 ↑	1.049	32	0.364	0.886	1.614
Controllabilità	1.461 ↑	1.053	32	0.365	1.096	1.826
Stimolazione	1.008 ↑	1.156	32	0.401	0.607	1.408
Originalità	0.359 →	1.160	32	0.402	-0.043	0.761

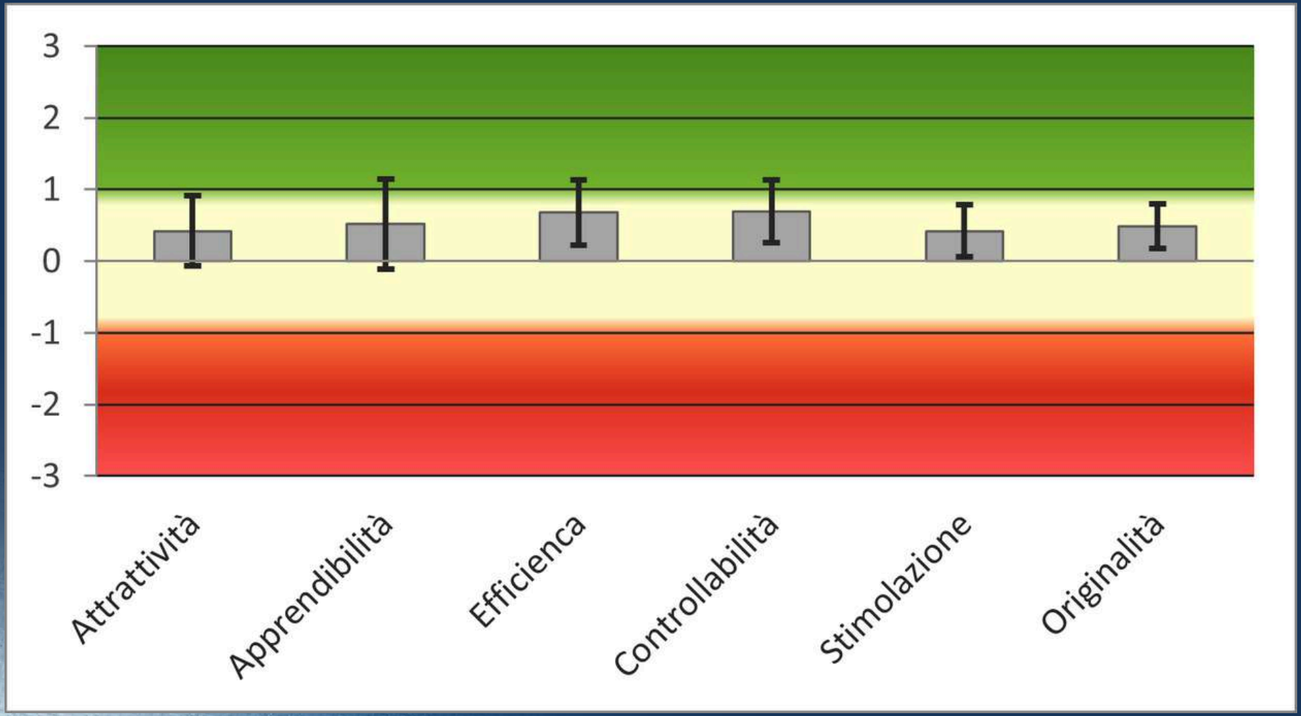
↑ Valori superiori a 0.8; indicano una **valutazione positiva**
→ Valori tra -0.8 e 0.8; indicano una **valutazione neutra**
↓ Valori inferiori a -0.8; indicano una **valutazione negativa**



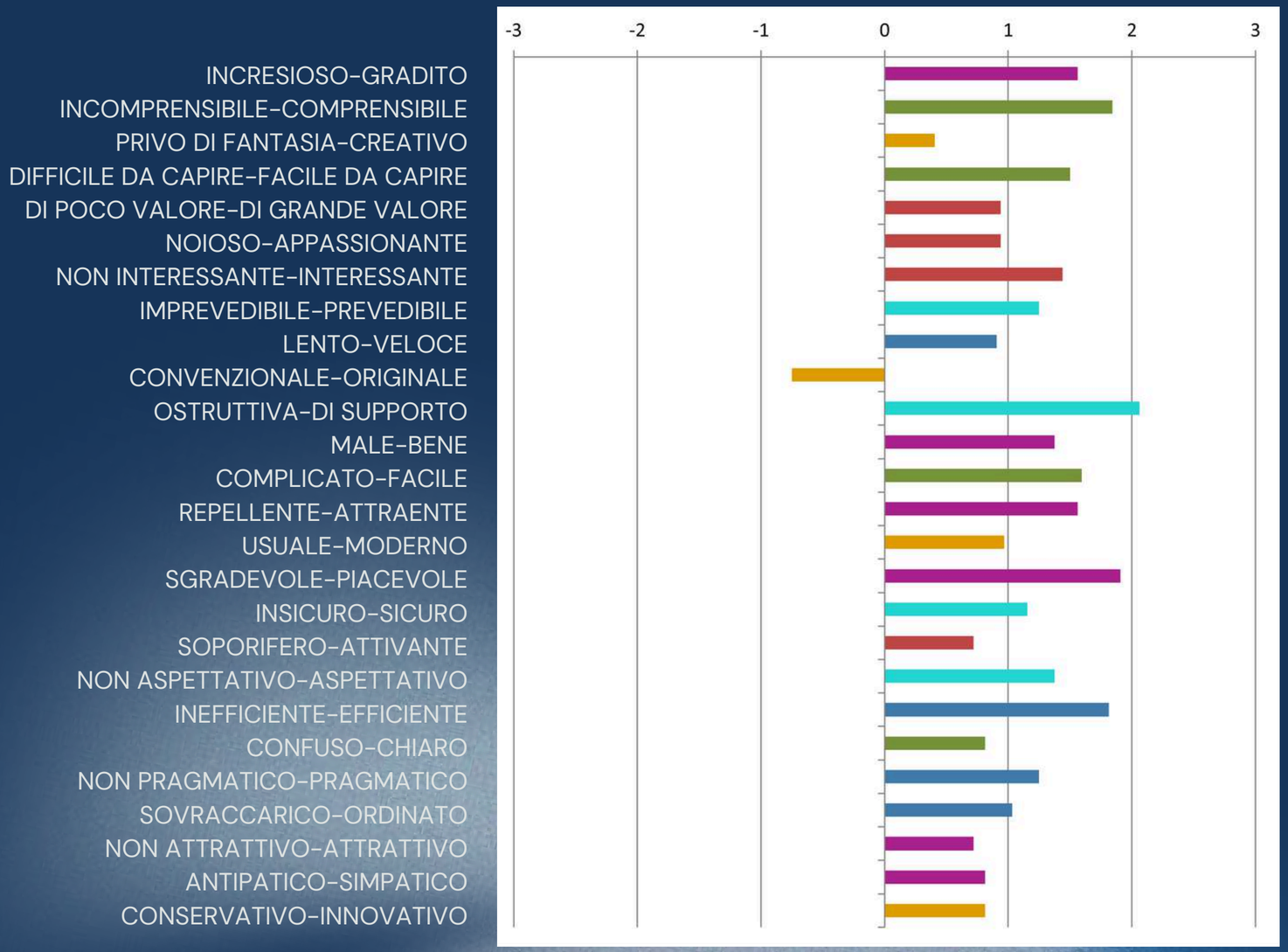
Analisi dei sottogruppi - Notability

Sottogruppo	Media	Dev. Standard	No.	Confidenza	I.C. (p = 0.05)	
Attrattività	0.422 →	1.412	32	0.489	-0.067	0.911
Apprendibilità	0.516 →	1.812	32	0.628	-0.112	1.143
Efficienca	0.680 →	1.319	32	0.457	0.223	1.137
Controllabilità	0.695 →	1.274	32	0.442	0.254	1.137
Stimolazione	0.422 →	1.042	32	0.361	0.061	0.783
Originalità	0.484 →	0.902	32	0.313	0.172	0.797

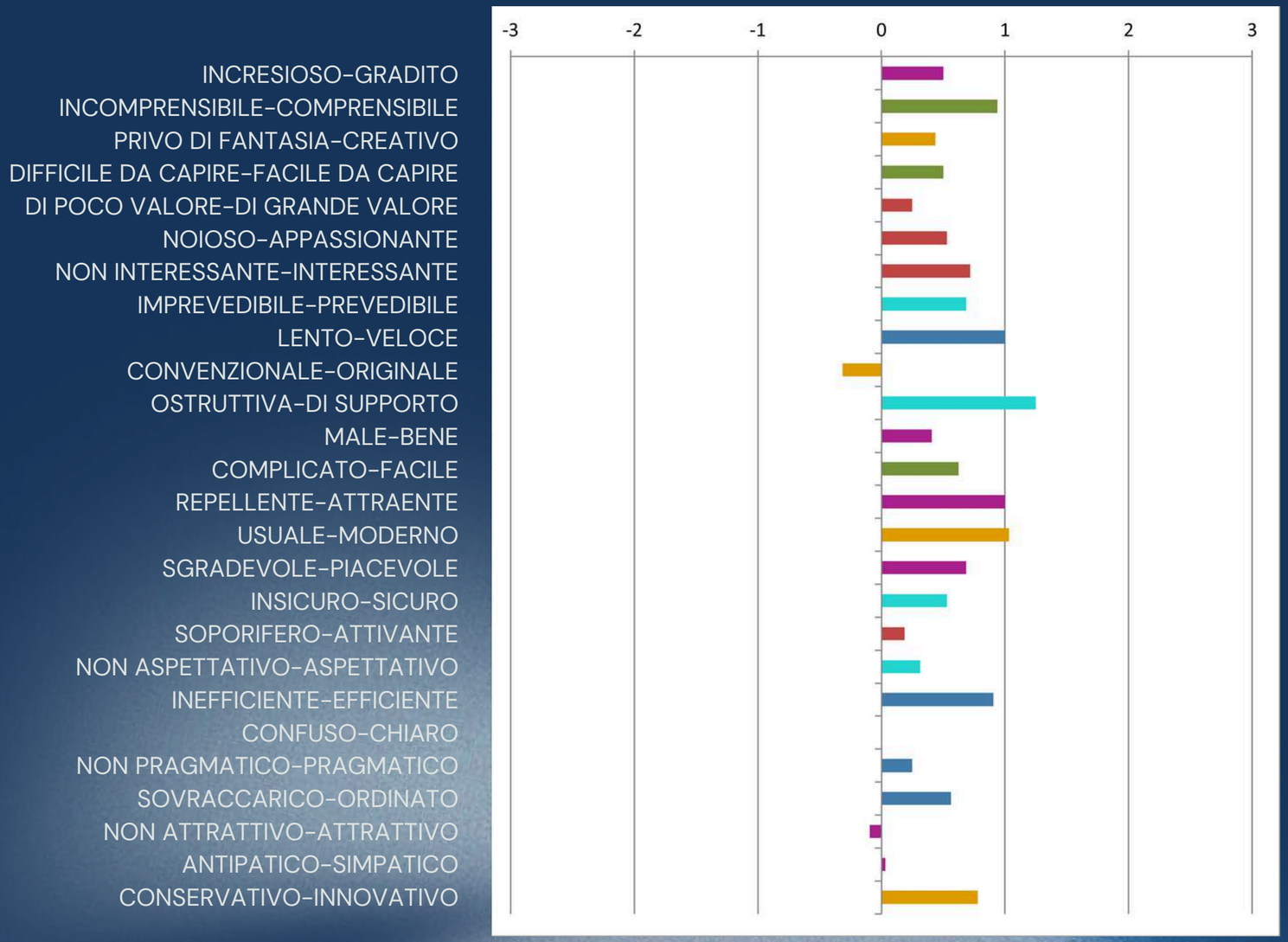
↑ Valori superiori a 0.8; indicano una **valutazione positiva**
→ Valori tra -0.8 e 0.8; indicano una **valutazione neutra**
↓ Valori inferiori a -0.8; indicano una **valutazione negativa**



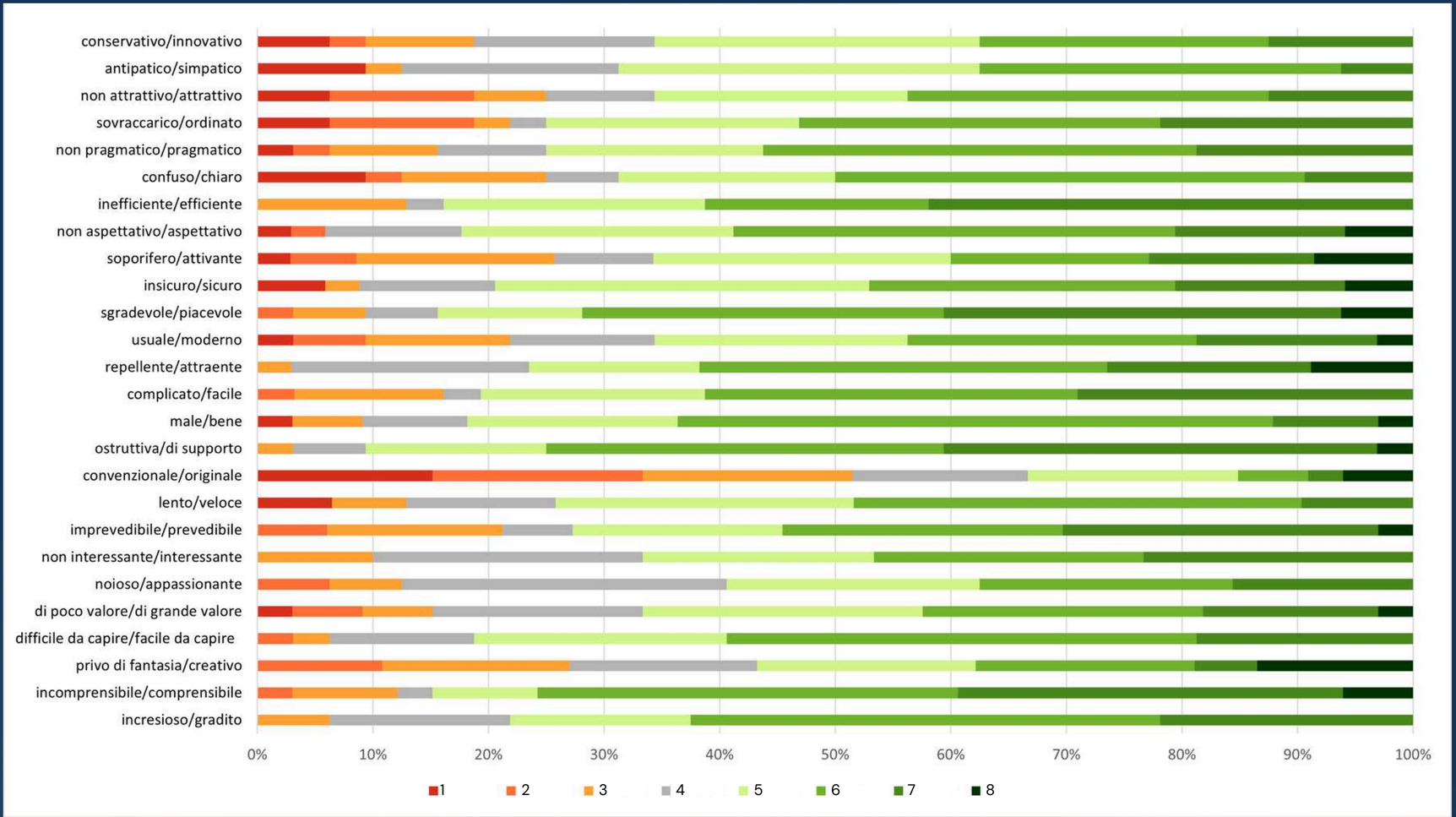
Valore medio delle risposte - GoodNotes



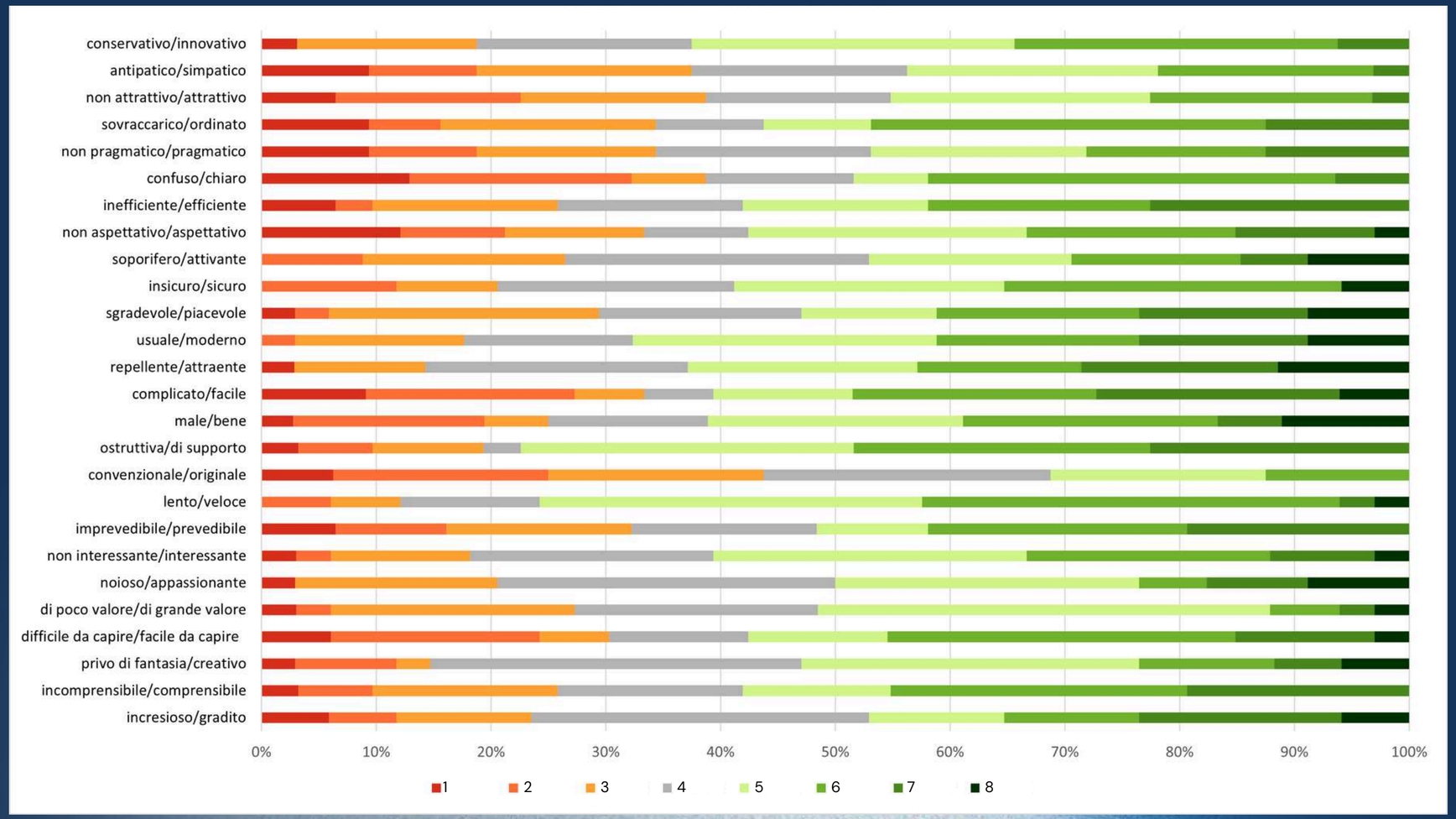
Valore medio delle risposte - Notability



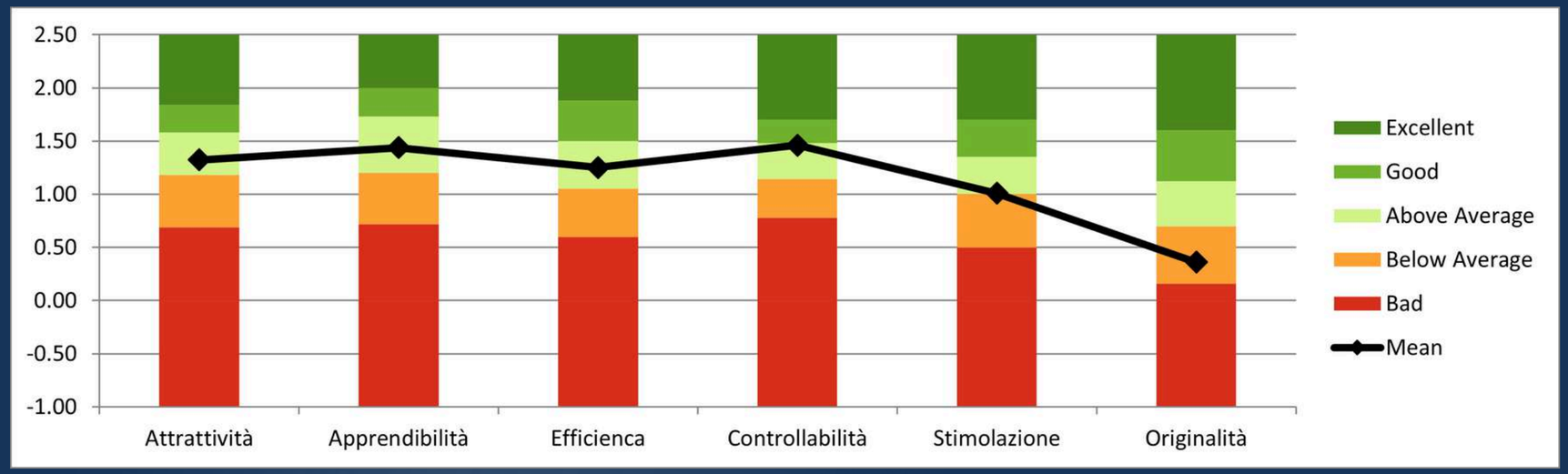
Distribuzione delle risposte - GoodNotes



Distribuzione delle risposte - Notability

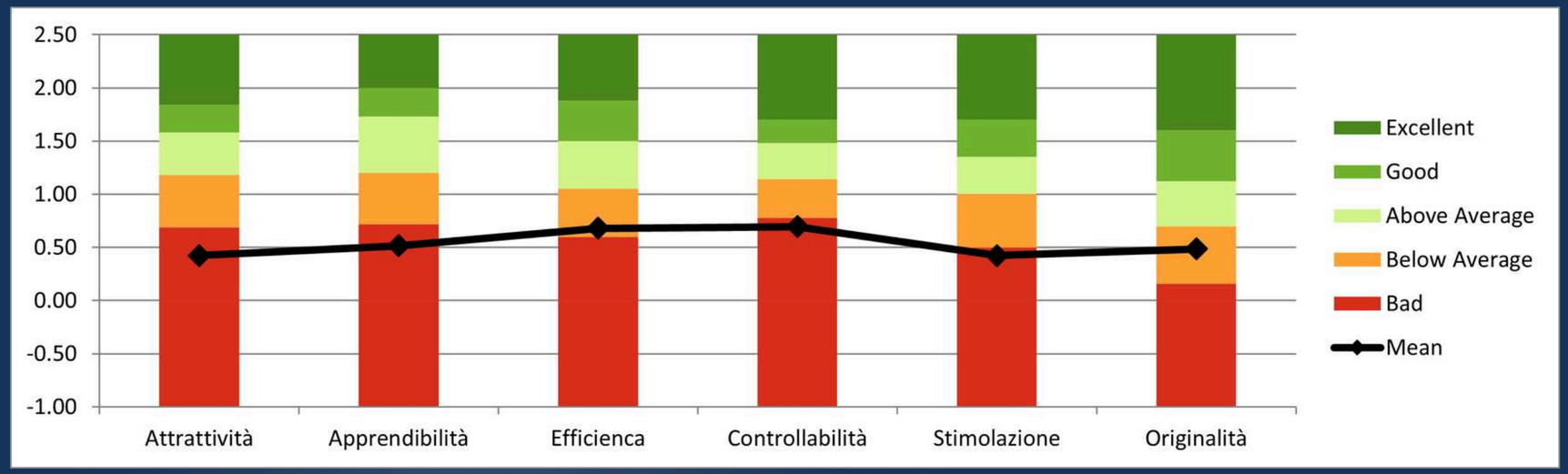


Confronto con il benchmark - GoodNotes



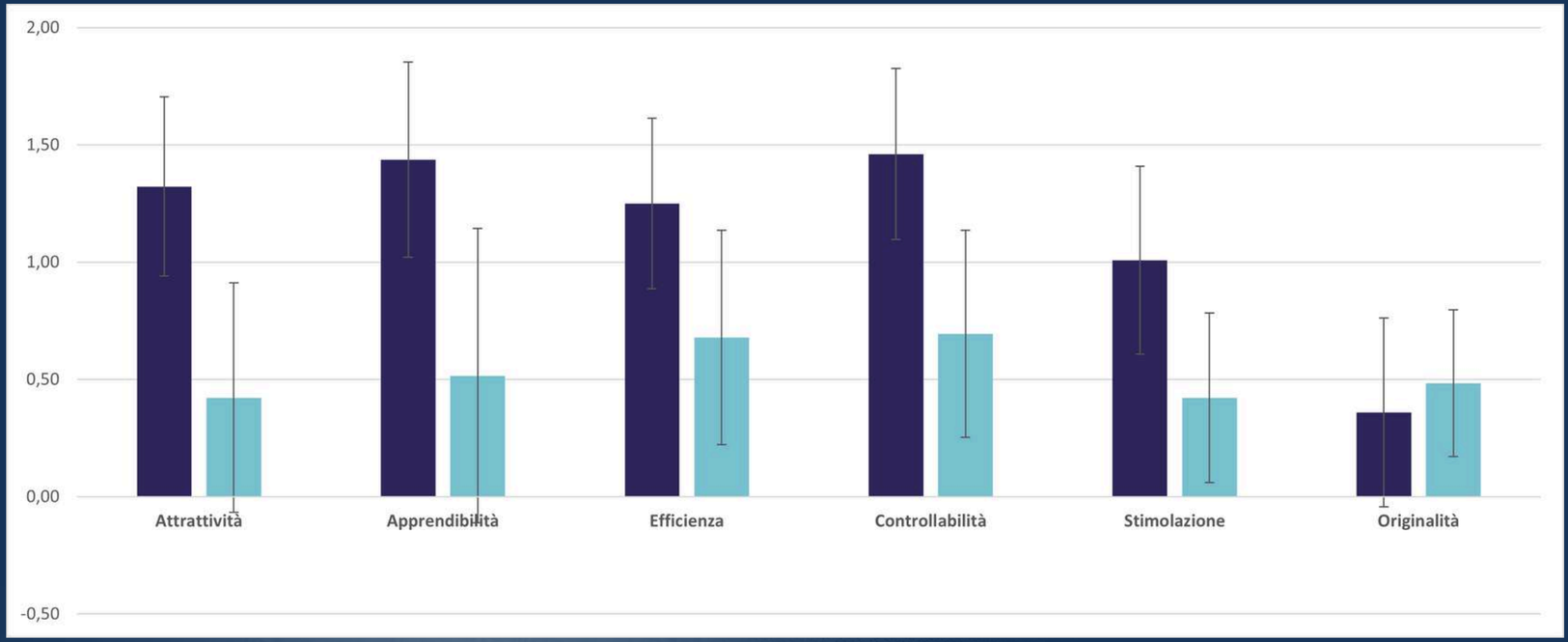
SOTTOGRUPPO	MEDIA	CONFRONTO	INTERPRETAZIONE
Attrattività	1.32	Sopra la media	Il 25% dei risultati fa meglio, il 50% peggio
Apprendibilità	1.44	Sopra la media	Il 25% dei risultati fa meglio, il 50% peggio
Efficienza	1.25	Sopra la media	Il 25% dei risultati fa meglio, il 50% peggio
Controllabilità	1.46	Sopra la media	Il 25% dei risultati fa meglio, il 50% peggio
Stimolazione	1.01	Sopra la media	Il 25% dei risultati fa meglio, il 50% peggio
Originalità	0.36	Sotto la media	Il 50% dei risultati fa meglio, il 25% peggio

Confronto con il benchmark - Notability



SOTTOGRUPPO	MEDIA	CONFRONTO	INTERPRETAZIONE
Attrattività	0,42	Scarso	Nell'intervallo del 25% dei peggiori risultati
Apprendibilità	0,52	Scarso	Nell'intervallo del 25% dei peggiori risultati
Efficienza	0,68	Sotto la media	Il 50% dei risultati fa meglio, il 25% peggio
Controllabilità	0,70	Scarso	Nell'intervallo del 25% dei peggiori risultati
Stimolazione	0,42	Scarso	Nell'intervallo del 25% dei peggiori risultati
Originalità	0,48	Sotto la media	Il 50% dei risultati fa meglio, il 25% peggio

Confronto dei valori medi



Sottogruppo	GOODNOTES						NOTABILITY						p-value	Differenza (livello 95%)
	Media	Dev. Standard	No.	Confidenza	I.C.		Media	Dev. Standard	No.	Confidenza	I.C.			
Attrattività	1,32	1,10	32	0,38	0,94	1,70	0,42	1,41	32	0,49	-0,07	0,91	0,0061	Significativa
Apprendibilità	1,44	1,20	32	0,42	1,02	1,85	0,52	1,81	32	0,63	-0,11	1,14	0,0199	Significativa
Efficienza	1,25	1,05	32	0,36	0,89	1,61	0,68	1,32	32	0,46	0,22	1,14	0,0604	Non significativa
Controllabilità	1,46	1,05	32	0,36	1,10	1,83	0,70	1,27	32	0,44	0,25	1,14	0,0111	Significativa
Stimolazione	1,01	1,16	32	0,40	0,61	1,41	0,42	1,04	32	0,36	0,06	0,78	0,0372	Significativa
Originalità	0,36	1,16	32	0,40	-0,04	0,76	0,48	0,90	32	0,31	0,17	0,80	0,6323	Non significativa

Net Promoter Score (NPS)

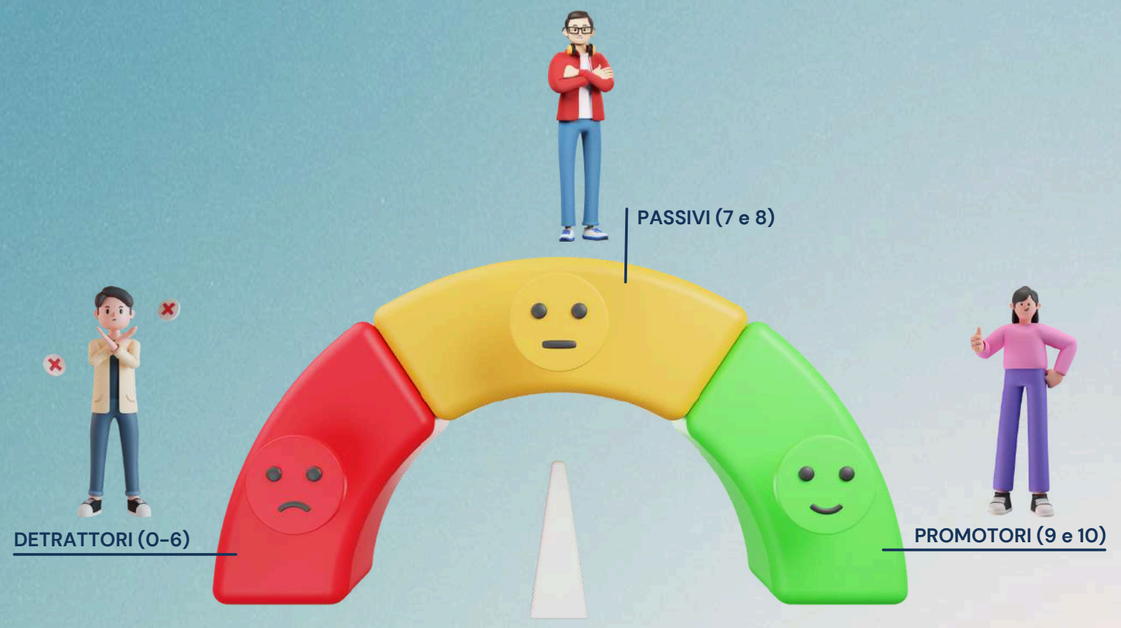
Il Net Promoter Score è un indice che permette di **valutare il livello di fedeltà e di apprezzamento** dei clienti nei confronti di un determinato servizio.

L'NPS si calcola sulle risposte date dagli utenti su una scala da 0 a 10 alla domanda:
"Con quale probabilità consiglieresti [nome del servizio] ad un amico o ad un collega?"

Le risposte vengono classificate in 3 fasce:

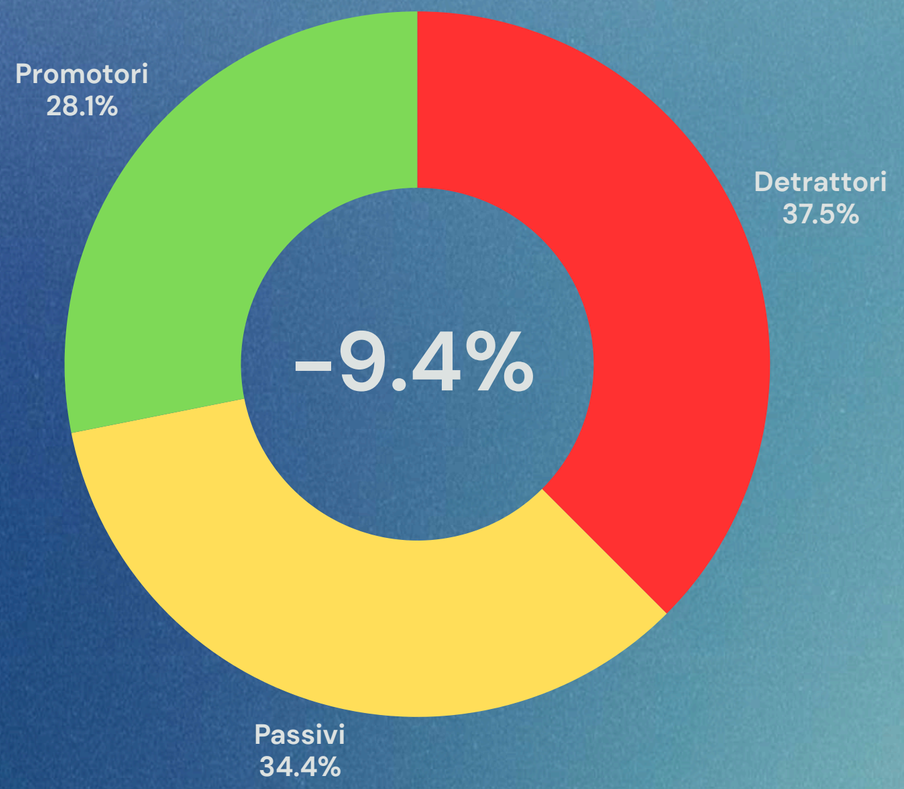
- **Detrattori** (da 0 a 6): utenti insoddisfatti.
- **Passivi** (7 e 8): utenti soddisfatti ma indifferenti; non consiglieranno né sconsiglieranno il prodotto.
- **Promotori** (9 e 10): utenti pienamente soddisfatti che consiglierebbero sicuramente il servizio ad altre persone.

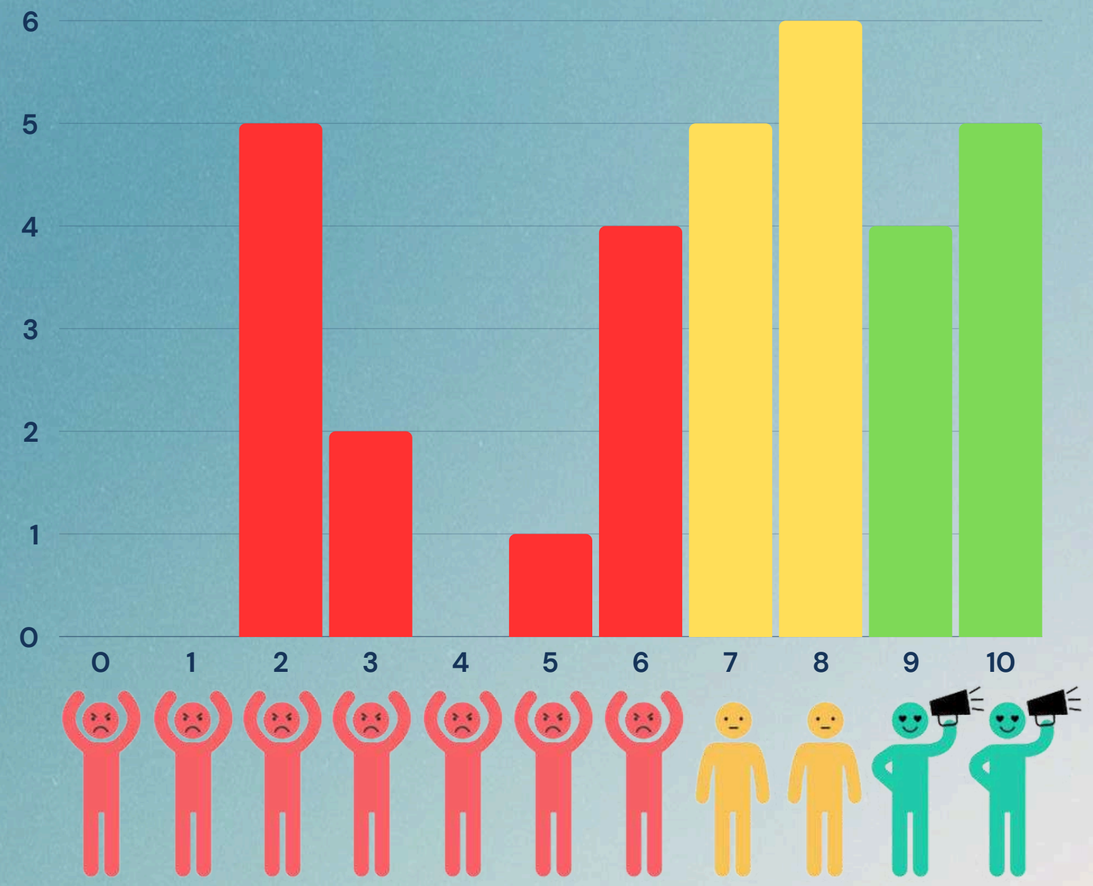
Dopo aver raccolto le risposte l'NPS viene calcolato **sottraendo la percentuale di detrattori da quella dei promotori**, ottenendo così un valore compreso tra -100 (tutti detrattori) e +100 (tutti promotori).



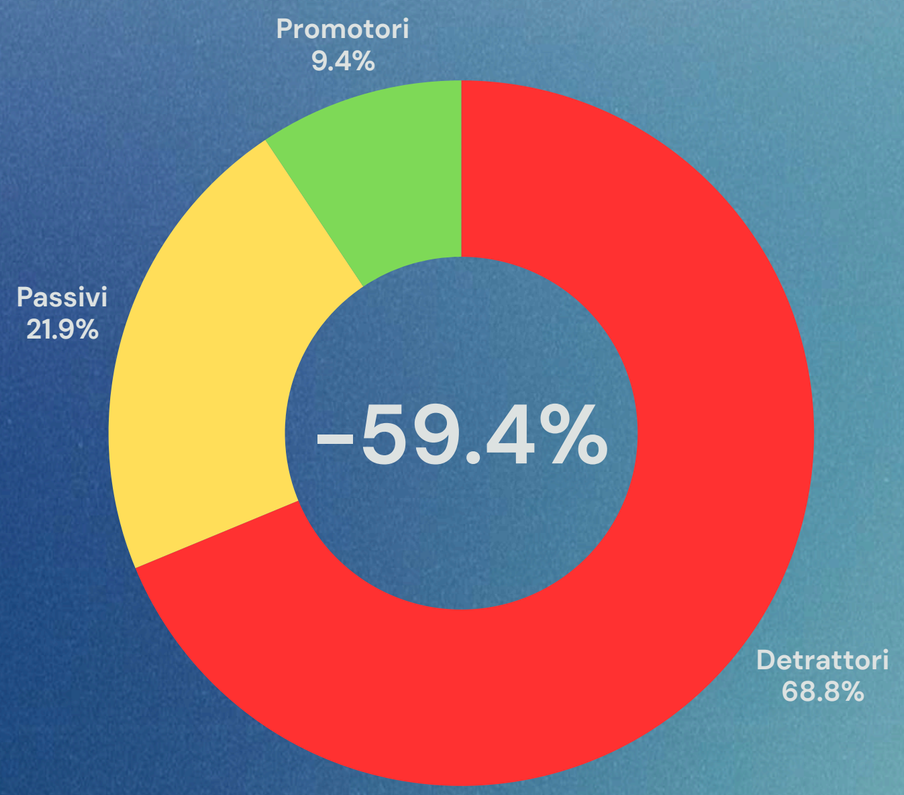
NPS = %PROMOTORI – %DETRATTORI

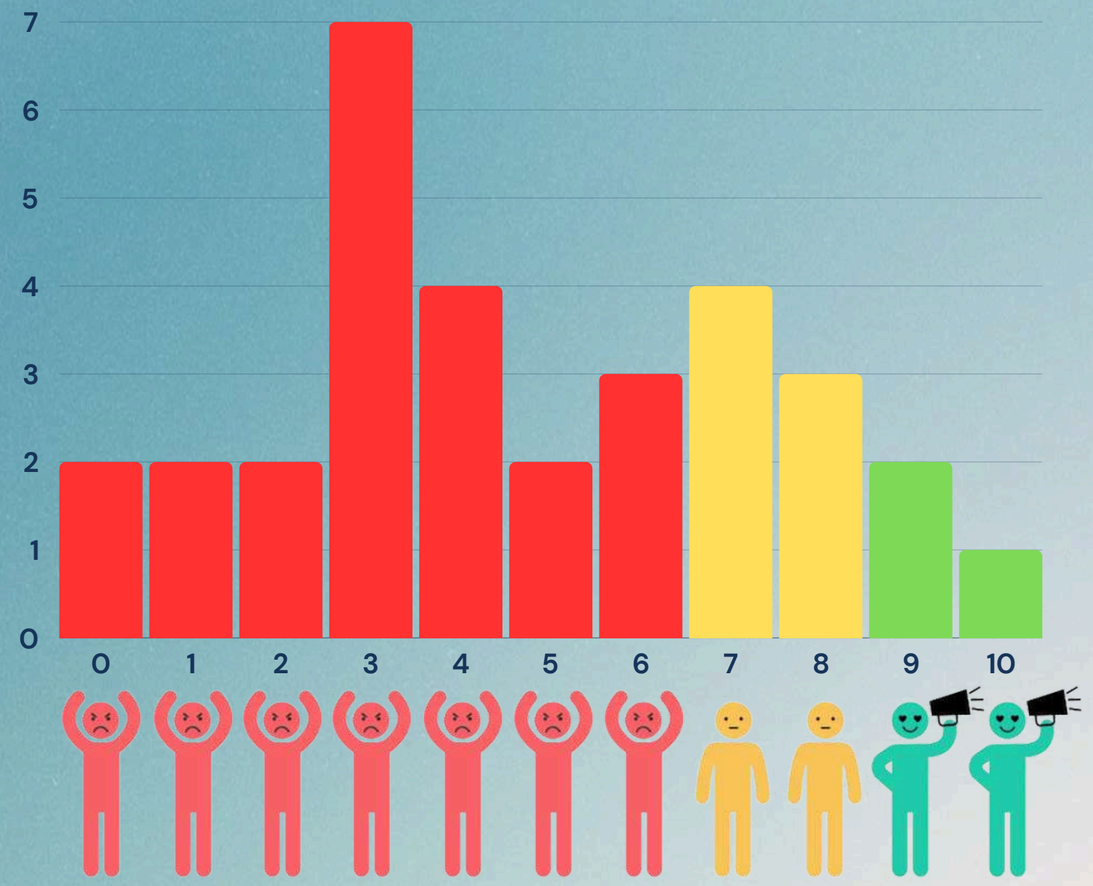
NPS GoodNotes





NPS Notability





QUESTIONARIO:

Considerazioni Finali

1

RISULTATI UEQ:
GoodNotes risulta essere più efficiente e piacevole rispetto a Notability. Nei sottogruppi del questionario UEQ, GoodNotes è significativamente migliore in attrattività, apprendibilità, controllabilità e stimolazione. Al contrario, Notability non risulta migliore in nessuna categoria.

2

CONFRONTO CON IL BENCHMARK:
Per quanto riguarda il benchmark, Goodnotes si posiziona sopra la media in tutte le categorie tranne stimolazione, che è nella media, e originalità, che è sotto la media. Notability, invece, ha punteggi molto negativi in tutte le categorie tranne efficienza e originalità, che sono leggermente sotto la media.

3

RISULTATI NPS:
Il punteggio NPS di Goodnotes è -9.4, indicativo di un feedback neutro. Al contrario, Notability ha uno score NPS di -59.4, riflettendo un feedback nettamente negativo.

CONCLUSIONI

Nonostante le due applicazioni apparissero inizialmente molto simili per approccio, pubblico di riferimento e finalità, un'analisi approfondita rivela che il loro indice di usabilità è diverso.

VS

- **Durante la prima fase, sono stati trovati 21 problemi di usabilità su Goodnotes, 4 in più di Notability.**
- **Inoltre, i valutatori hanno dichiarato di percepire Goodnotes come più organizzata e intuitiva rispetto a Notability, che è stata considerata più complessa e dispersiva, nonostante alcune scelte di design di Goodnotes siano state ritenute ambigue.**
- **Successivamente, durante i test con gli utenti, sono emerse alcune differenze tra l'efficacia e l'efficienza dei due sistemi:**
  - **In particolare, per quanto riguarda il terzo task di Notability, è stato riscontrato un tasso di errori superiore e un tempo di esecuzione maggiore rispetto a Goodnotes.**
- **Nell'indagine sulla soddisfazione degli utenti, il confronto con il benchmark ha evidenziato l'eccellenza di Goodnotes in aspetti come attrattività, apprendibilità, efficienza e controllabilità, mentre Notability è risultato sotto la media in questi ambiti.**

APPENDICE

Link per consultare le risorse utilizzate:

https://bit.ly/GoodNotes-VS-Notability

QR Code

GRAZIE

De Marchi Sara - Ferrara Luigi - Leskovec Ethan - Romaniello Andrea