

UNIVERSITÀ DEGLI STUDI DI TORINO

Dipartimento di Culture, Politica e Società

Corso di Laurea Magistrale in Comunicazione, ICT e Media

Classe di Laurea LM-59



Tesi di Laurea Magistrale in

Interazione Uomo-Macchina: Approcci Avanzati

Come l'empatia trasforma l'interazione uomo-robot

Uno studio condotto con il robot Pepper per valutare le sue capacità di interazione e persuasione

Relatrice:

Prof.ssa Cristina Gena

Correlatrice:

Prof.ssa Fabiana Venero

Candidata:

Giorgia Buracchio

Anno accademico 2023-2024

SOMMARIO

1. Introduzione	4
2. Stato dell'arte	6
2.1 Introduzione alla Robotica Sociale e all'Interazione Uomo-Robot	6
2.2 Comunicazione empatica nei sistemi robotici	7
2.2.1 Emozioni e affective computing	8
2.2.2 Modello di Plutchik delle emozioni	10
2.3 La Persuasione nella robotica sociale	12
2.3.1 Empatia a persuasione: un caso di studio	14
2.3.2 Personalità e persuasione: un caso di studio	16
2.4 Il robot Pepper	18
2.4.1 Caso di studio 1 con Pepper	19
2.4.2 Caso di studio 2 con Pepper	20
2.5 Riflessioni	21
3. Progettazione, implementazione e verifica del dialogo	23
3.1 Tema dell'esperimento e supporto scientifico	23
3.1.1 Considerazioni etiche e sociali	24
3.2 Prompt engineering	25
3.2.1 Pre Generazione dei prompt per robot empatico	26
3.2.2 Large Language Models e il modello COSTAR	26
3.2.3 La tecnica "One/Few Shot"	29
3.3 Settaggio di Pepper	31
3.3.1 I traduttori confrontati e analisi	31
3.3.2 Implementazione del dialogo	33
3.3.2.1 Dialogo empatico	34
3.3.2.2 Dialogo neutro	35
3.3.3 Implementazione immagini	35
3.3.3.1 Dialogo empatico	36

3.3.3.2 Dialogo neutro	37
3.4 Verifica del sistema empatico	38
3.4.1 Test digitale	40
3.4.2 Interviste di test	41
4. Sperimentazione e analisi	43
4.1 Disegno sperimentale	43
4.1.1 Ipotesi sperimentale	44
4.1.2 Partecipanti	45
4.1.3 Ambiente e materiali	46
4.2 Procedura sperimentale	50
4.2.1 Raccolta dati	53
4.3 Analisi dati	54
4.3.1 Analisi preliminare	54
4.3.2 Kappa di Fleiss	56
4.3.3 Misurazione dell'empatia di Pepper	61
4.4 Risultati	62
4.4.1 Persuasività	63
4.4.2 Stile comunicativo	64
4.4.3 Analisi dell'Auto-Rivelazione	65
5. Conclusioni	67
5.1 Limiti e vizi sperimentali	67
5.2 Scenari futuri	69
6. Riferimenti bibliografici e digitali	71
6.1 Bibliografia	71
6.2 Sitografia	77

1. Introduzione

L'interazione uomo-robot è un campo di ricerca multidisciplinare che coinvolge vari ambiti come la robotica, l'intelligenza artificiale, la psicologia e le scienze cognitive. L'obiettivo principale è sviluppare robot capaci di interagire con gli esseri umani in modo naturale e intuitivo, creando una comunicazione bidirezionale che sia efficace e significativa. La robotica sociale, una sottodisciplina della robotica, si concentra su aspetti come l'empatia, la comprensione delle emozioni e la capacità dei robot di rispondere in modo adeguato agli stati emotivi degli utenti. In questo contesto, la capacità di interpretare e reagire alle emozioni umane non è solo una caratteristica desiderabile, ma diventa cruciale per l'interazione empatica, un elemento fondamentale per stabilire una connessione significativa tra uomo e robot.

La comprensione delle emozioni rappresenta uno dei pilastri su cui si fonda l'interazione empatica. Un robot in grado di riconoscere gli stati emotivi delle persone e rispondere in maniera appropriata può facilitare il dialogo e migliorare l'esperienza interattiva, rendendola più coinvolgente e personale. Questa capacità diventa particolarmente importante in contesti come l'assistenza, l'educazione e la compagnia, dove il robot non è solo uno strumento tecnologico, ma un'entità con la quale si stabilisce una relazione. L'obiettivo di creare robot capaci di comprendere e rispondere alle emozioni umane è quindi un passo avanti verso una comunicazione più profonda e un'interazione sociale più autentica.

La tesi si inserisce in questo ampio quadro di ricerca, partendo da un progetto denominato "don't cHRI", condotto presso l'Advanced Interaction and Robotics Lab del Dipartimento di Informatica dell'Università degli Studi di Torino. Questo progetto ha come obiettivo l'esplorazione degli effetti dell'empatia generata e mostrata dal robot sociale Pepper sugli studenti universitari. Nello studio condotto, particolare attenzione è stata posta sull'influenza che il comportamento empatico di Pepper può avere sull'accettazione e sull'attaccamento dei partecipanti a determinati temi di conversazione. La ricerca, si concentra dunque su come le caratteristiche empatiche del robot possano modificare la percezione e l'interazione degli utenti con il robot stesso, influenzando non solo l'esperienza di comunicazione, ma anche le opinioni e le emozioni dei partecipanti.

L'interazione empatica del robot non si limita alla semplice risposta verbale, ma coinvolge una serie di comportamenti multimodali che includono espressioni non verbali, gesti e, in alcuni casi, la presentazione di contenuti visivi. Questo approccio, combinato con tecniche avanzate di "prompt engineering", permette di stimolare reazioni emotive specifiche negli utenti. Il progetto "don't cHRI" ha offerto un'opportunità per esplorare queste dinamiche, fornendo una base solida per testare le ipotesi legate all'efficacia persuasiva del robot e al modo in cui esso può favorire una comunicazione più aperta e personale.

Lo studio si colloca quindi nel contesto della robotica sociale, con l'obiettivo di esaminare in che modo l'empatia artificiale possa influenzare l'interazione umana. La ricerca affronta anche domande fondamentali su come il comportamento del robot possa influire sulle emozioni e sui comportamenti degli utenti, contribuendo a delineare nuove direzioni per l'interazione uomo-robot. In particolare, l'esperimento condotto ha permesso di raccogliere dati sull'impatto del comportamento empatico del robot, fornendo un punto di partenza per ulteriori esplorazioni nel campo della robotica sociale e della comunicazione persuasiva mediata dalla tecnologia.

Concludendo, questa tesi si propone di analizzare i risultati emersi dall'esperimento condotto con Pepper, offrendo una riflessione su come l'interazione empatica possa essere utilizzata per promuovere una comunicazione più efficace e coinvolgente tra robot e utenti umani. Attraverso l'analisi di questi risultati, si esploreranno le implicazioni della robotica sociale non solo dal punto di vista tecnologico, ma anche dal punto di vista delle scienze cognitive e comportamentali, con particolare attenzione ai possibili scenari futuri e alle applicazioni dello stesso progetto su robot che hanno differenti caratteristiche soprattutto fisiche, come il robot NAO.

2. Stato dell'arte

2.1 Introduzione alla Robotica Sociale e all'Interazione Uomo-Robot

L'interazione uomo-robot (Human-Robot Interaction, HRI) è un campo di studio multidisciplinare che si concentra sulla progettazione, sviluppo e valutazione di sistemi robotici pensati per essere utilizzati o per collaborare con gli esseri umani. Questa disciplina si dedica a comprendere e migliorare le modalità di comunicazione tra esseri umani e robot, cercando di rendere queste interazioni più naturali ed efficaci.

La comunicazione tra esseri umani e robot può essere classificata in due categorie principali: interazioni remote e interazioni prossimali. Le interazioni remote comprendono la teleoperazione e il controllo di supervisione, dove l'operatore umano può gestire il robot da una posizione distante. Questo tipo di interazione è cruciale in scenari come le missioni di ricerca e salvataggio, dove l'ambiente può essere pericoloso per gli esseri umani. Le interazioni prossimali, invece, includono sia interazioni fisiche che interazioni sociali. (Goodrich & Shultz)

Queste ultime vengono adottate nella robotica sociale, un ramo della robotica dedicato allo sviluppo di robot progettati per interagire e comunicare con gli esseri umani su un livello sociale e personale. (Breazeal, 2003). La robotica sociale ha guadagnato importanza nel contesto moderno grazie al suo potenziale di migliorare la qualità della vita in molteplici settori, inclusi la sanità, l'istruzione e l'industria. Ad esempio, robot come Pepper, sviluppati da SoftBank Robotics, sono utilizzati in ambienti commerciali per assistere i clienti e fornire informazioni. La capacità dei robot sociali di interagire in modo empatico con gli esseri umani li rende strumenti preziosi per la terapia, l'assistenza agli anziani e l'educazione dei bambini (Dautenhahn, 2007).

Un aspetto centrale della HRI è la comunicazione, esplorata a diversi livelli, concentrandosi su comunicatore, messaggio e conversazione. La teoria della comunicazione applicata alla HRI si avvale del metamodello costitutivo come disciplina pratica per inquadrare le interazioni uomo-robot. Le linee di ricerca in HRI si distinguono in base alle tradizioni semiotica, sociopsicologica e socioculturale. È fondamentale collegare la teoria della comunicazione alla HRI per comprendere meglio le interazioni uomo-robot, utilizzando la teoria dell'interazione uomo-uomo come punto di riferimento (Frijns et al., 2023).

Lo stile di interazione di un robot svolge un ruolo cruciale nell'interazione uomo-robot, influenzando significativamente l'autoefficacia delle persone, la percezione del robot e l'impegno nei compiti. Le persone tendono a impegnarsi più a lungo nei compiti quando interagiscono con un robot che fornisce un feedback legato alla persona, evidenziando l'importanza di incorporare elementi sociali nelle interazioni con i robot. I robot che offrono stili di interazione orientati alla persona possono migliorare l'autoefficacia degli individui nell'HRI, portando a una maggiore accettazione percepita del robot e a una riduzione della frustrazione durante le interazioni (Zafari et al., 2019).

La ricerca sull'HRI si concentra sul miglioramento dell'interazione tra esseri umani e robot, studiando come i robot durante la comunicazione, possono comprendere e rispondere alle emozioni umane. Gli studi hanno dimostrato che una buona esperienza utente (user experience - UX) è fondamentale per l'accettazione e l'adozione dei robot sociali. Ad esempio, uno studio recente ha evidenziato che la valutazione dell'UX nei robot sociali è essenziale per garantire interazioni positive e per migliorare la fiducia e l'accettazione dei robot da parte degli utenti (Conti et al., 2017). Nello studio di Conti et al., è stato riscontrato che i robot sono strumenti utili ed efficaci nell'educazione e nella terapia dei bambini con disabilità dello sviluppo, sebbene siano emerse preoccupazioni riguardo alla riduzione dell'interazione umana e alla necessità di una formazione adeguata.

La capacità di garantire una UX positiva è strettamente legata alla comunicazione empatica, che è essenziale per creare interazioni significative e coinvolgenti tra robot e umani. La comunicazione empatica non solo migliora la UX, ma è anche fondamentale per la costruzione di relazioni di fiducia tra l'utente e il robot.

2.2 Comunicazione empatica nei sistemi robotici

L'empatia è fondamentale nella HRI, facilitando la formazione e il mantenimento di relazioni sociali con i robot. Tuttavia, sorgono domande sulla possibilità di applicare il concetto di empatia, profondamente radicato nei contesti sociali umani, ai robot (Malinowska, 2021). Lo sviluppo di agenti virtuali e robot dotati di empatia computazionale può portare a interazioni più credibili e naturali con gli utenti in vari contesti. Incorporando l'empatia nelle macchine, potrebbe essere possibile influenzare positivamente i comportamenti umani, promuovendo risposte prosociali sia negli esseri umani che nelle macchine. Gli esseri umani tendono a

imitare inconsciamente i comportamenti degli altri, seguire le espressioni e rispondere in modo empatico, portando a interazioni più fluide e a un maggiore legame. Questo dimostra come l'empatia possa influenzare positivamente il comportamento umano, favorendo le interazioni sociali e portando a comportamenti di cura e aiuto. Comprendere e replicare questi meccanismi nei modelli computazionali potrebbe, quindi, modellare potenzialmente i comportamenti umani, rendendo le interazioni con le macchine più naturali e significative (Paiva et al., 2017).

2.2.1 Emozioni e affective computing

L'integrazione delle emozioni e degli stati d'animo nei robot è essenziale per garantire interazioni fluide e naturali. Le emozioni espressive e gli stati d'animo sono fondamentali nella trasmissione di messaggi e nella costruzione di relazioni a lungo termine tra esseri umani e robot. Il modello affettivo sviluppato per i robot sociali si concentra sulla creazione di affetti naturali e simili a quelli umani, distinguendo tra risposte emotive immediate, stato d'animo generale e atteggiamenti a lungo termine nei confronti di ciascun visitatore. Questo modello, illustrato nel contesto del robot *Roboceptionist* Valerie (vedi figura 1), ha dimostrato di migliorare le relazioni uomo-robot, evidenziando il potenziale per migliorare l'interazione e la comunicazione in vari ambiti come il servizio clienti, la sanità e l'istruzione (Kirby et al., 2010).



Figura 1: Le quattro espressioni utilizzate per testare l'effetto dell'umore del robot: positivo, neutro, negativo e triste (da sinistra a destra) durante l'interazione. Le differenze includono la posizione delle sopracciglia e delle palpebre, la forma della bocca e l'angolo della testa.

Fonte: <https://onlinelibrary.wiley.com/doi/full/10.1609/aimag.v32i4.2383>

L'affective computing (informatica affettiva), è un concetto cruciale in questo contesto. Rosalind Picard, pioniera dell'affective computing, lo definisce come “lo studio e lo sviluppo

di sistemi e dispositivi che possono riconoscere, interpretare, elaborare e simulare le emozioni umane” (Picard, 1997).

Incorporare emozioni e stati d’animo umani nei robot può portare a numerosi vantaggi. La comunicazione ne risulta migliorata, poiché i robot possono esprimere sentimenti come felicità, tristezza o empatia, rendendo le interazioni più coinvolgenti e relazionabili. Questo, a sua volta, porta a un’esperienza utente più positiva e a una maggiore soddisfazione degli utenti. I robot emotivamente espressivi sono in grado di comprendere lo stato emotivo degli utenti e rispondere di conseguenza, creando un ambiente di interazione più solidale. Inoltre, le emozioni simili a quelle umane nei robot contribuiscono a creare fiducia negli utenti, migliorando la cooperazione e l’accettazione dei robot stessi. Imitando le emozioni umane, i robot possono stabilire relazioni a lungo termine con gli utenti, favorendo un senso di connessione e compagnia (Kirby et al., 2010).

Il sistema comportamentale dei robot, piuttosto che il sistema emozionale, domina l’arbitraggio del comportamento, permettendo ai robot di impegnarsi in vari comportamenti legati al gioco con gli individui, migliorando le interazioni uomo-robot. Trattando i robot come creature socialmente reattive e condividendo i significati durante le interazioni, gli individui possono rendere il comportamento del robot più coerente e prevedibile, portando a esperienze di comunicazione e apprendimento più efficaci. Robot come Kismet (vedi figura 2), sviluppato da Cynthia Breazeal per il suo dottorato presso il MIT AI Lab alla fine degli anni Novanta, possono imparare dagli scambi sociali ed esplorare il ruolo delle emozioni nel processo di apprendimento. Kismet è il primo robot progettato per coinvolgere esplicitamente le persone in un’interazione naturale ed espressiva faccia a faccia, ispirato allo sviluppo sociale infantile, alla psicologia e all’etologia (Breazeal, 2000).

Queste interazioni, fondamentali per lo sviluppo sociale, emotivo e cognitivo, possono essere sfruttate per avviare lo sviluppo socio-emotivo-cognitivo dei robot sociali, migliorando in definitiva la loro funzionalità e usabilità in ambienti complessi e mutevoli (Breazeal, 2003).

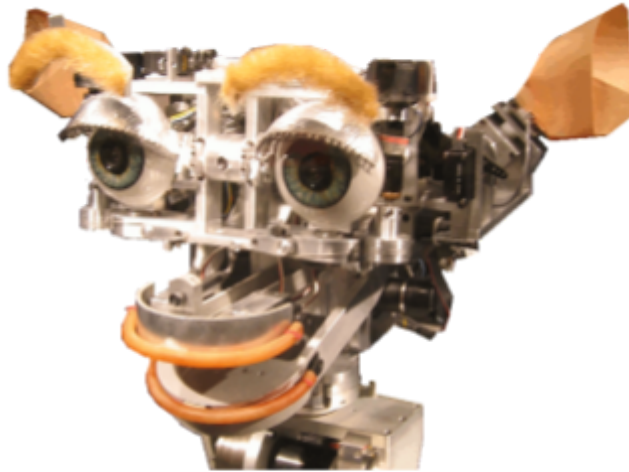


Figura 2: Kismet è in grado di generare una gamma continua di espressioni di varia intensità mescolando le posture facciali di base. Espressione mostrata è la felicità, includendo il movimento della bocca, delle ciglia, delle sopracciglia e delle orecchie.

Fonte: https://commons.wikimedia.org/wiki/File:Kismet_robot_20051016.gif

L'emozione è cruciale nel comunicare il proprio stato emotivo agli altri attraverso caratteristiche espressive della voce, del viso, dei gesti e della postura, influenzando il comportamento altrui. L'attivazione delle emozioni in robot come Kismet comporta espressioni facciali e comportamenti che servono a funzioni specifiche, come il comportamento di fuga innescato da stimoli minacciosi. Il coordinamento dei sistemi di riconoscimento affettivo, di emozione, di motivazione e di espressione nei robot influenza il loro comportamento emotivo complessivo durante l'impegno sociale con gli esseri umani, migliorando la qualità dell'interazione (Breazeal, 2003).

2.2.2 Modello di Plutchik delle emozioni

Robert Plutchik, psicologo americano di grande rilievo, è noto per il suo contributo alla teoria delle emozioni. Plutchik ha sviluppato la "ruota delle emozioni" (vedi figura 3), un modello che illustra le diverse emozioni umane e le loro interrelazioni.

Questo modello offre la possibilità di avere una classificazione affettiva più ricca e sfumata, poiché fornisce un modo operativo per classificare sia le emozioni primarie (gioia, fiducia,

paura, sorpresa, tristezza, disgusto, rabbia e anticipazione) che quelle complesse (chiamate anche diadi), derivanti dalla combinazione delle emozioni di base.



Figura 3: La “Ruota delle Emozioni” di Plutchik rappresenta le emozioni primarie e le loro interrelazioni. Le emozioni aumentano di intensità dal centro verso l’esterno e si combinano per formare emozioni complesse. La variazione di colore e la struttura del cono illustrano questa intensità, ad esempio, la rabbia varia da irritazione a collera e la tristezza da pensierosità ad angoscia.

Fonte: <https://italia.6seconds.org/2017/05/modello-delle-emozioni-plutchik-aggiornamenti-2017/>

Il modello di Plutchik offre anche un modo per rappresentare, attraverso la sua struttura spaziale, le emozioni e le loro interconnessioni. In questo modello, la distanza affettiva tra diversi stati emotivi è una funzione della loro distanza radiale. La dimensione verticale del cono rappresenta l’intensità: le emozioni si intensificano man mano che si muovono dall’esterno verso il centro della ruota, e questo è indicato anche dal colore: più scura è la tonalità, più intensa è l’emozione. Ad esempio, la rabbia al suo minimo livello di intensità è fastidio, mentre al suo massimo livello diventa collera. Analogamente, un sentimento di noia può intensificarsi fino a diventare disgusto se non viene considerato, rappresentato da una tonalità di viola scuro (*La Ruota Delle Emozioni Di Plutchik*, 2020).

La teoria di Plutchik è spesso confrontata con quella di Paul Ekman (Ekman & Keltner, 1970), che identifica sei emozioni base (gioia, tristezza, paura, rabbia, sorpresa e disgusto) che sono universali e riconoscibili attraverso espressioni facciali. Mentre Ekman si concentra

sulle espressioni facciali come indicatori delle emozioni, Plutchik offre una visione più dinamica e tridimensionale delle emozioni, mostrando come esse interagiscono e si trasformano. L'ontologia di Plutchik codifica le categorie emotive in una tassonomia che rappresenta:

- le emozioni di base o primarie;
- le emozioni complesse (o composte);
- l'opposizione tra le emozioni;
- la somiglianza tra le emozioni.

In conclusione, la teoria psico evolutiva generale delle emozioni di Plutchik, si distingue per la sua complessità e il suo valore andando a arricchire la nostra comprensione delle dinamiche emotive. Essa fornisce preziose indicazioni su come le emozioni si relazionano con l'adattamento e i processi evolutivi, evidenziando la loro rilevanza sia per la psicologia evolutiva che per la ricerca contemporanea (Plutchik, 1980).

2.3 La Persuasione nella robotica sociale

La persuasione è un meccanismo attraverso cui si cerca di influenzare le decisioni e i comportamenti degli altri, usando diverse tecniche e strategie per ottenere un determinato risultato. Nella robotica sociale, questo concetto, viene applicato in modo innovativo con i robot progettati per interagire con gli esseri umani in modo convincente e mirato, al fine di influenzare comportamenti e decisioni in maniera efficace (Fiore, 2019).

Robert Cialdini, psicologo e ricercatore dell'Università dell'Arizona, ha approfondito il tema della persuasione attraverso il suo libro "The Science of Persuasion" (Cialdini, 2001). In questo lavoro, Cialdini esplora i meccanismi che ci portano a dire di sì, identificando sei categorie principali di tecniche persuasive.

Una delle tecniche principali è la reciprocità, che si basa sul principio di dare e ricevere. Quando qualcuno riceve qualcosa, sente un obbligo di ricambiare il favore. Questo fenomeno è così radicato nella nostra psicologia che ignorarlo può portare a essere etichettati come ingrati. Ad esempio, offrire un campione gratuito di un prodotto spesso aumenta la probabilità

che il destinatario acquisti il prodotto stesso, poiché si sente in dovere di contraccambiare il gesto.

Un'altra tecnica è l'impegno e la coerenza. Le persone tendono a mantenere un'immagine coerente nel tempo per evitare di essere viste come inaffidabili o superficiali. Questo principio sfrutta la nostra avversione al disallineamento tra ciò che affermiamo e ciò che facciamo. Ad esempio, un cambiamento improvviso nella *brand image* può confondere i consumatori, che tendono a preferire la coerenza.

La riprova sociale è particolarmente rilevante. Questo principio psicologico, descritto da Cialdini, si basa sull'idea che le persone tendono a conformarsi al comportamento degli altri, specialmente quando non sono sicure di quale sia il comportamento appropriato in una determinata situazione (*Il Principio Della Riprova Sociale Secondo Robert Cialdini*, n.d.).

Nel contesto della robotica sociale, la riprova sociale può essere utilizzata per guidare le interazioni tra robot e umani. Per esempio, un robot progettato per promuovere comportamenti salutarì potrebbe mostrare che altri utenti, o "amici" robotici, stanno già adottando tali comportamenti. Questo potrebbe influenzare l'utente umano a seguire l'esempio, basandosi sull'assunto che se altre persone stanno agendo in un certo modo, è probabile che sia la scelta giusta.

Per valutare l'efficacia delle strategie persuasive dei robot sociali, è cruciale fare affidamento su misure oggettive, come il comportamento reale dei partecipanti, piuttosto che su misure soggettive come la loro percezione del robot. Le osservazioni concrete dell'interazione e dei cambiamenti comportamentali forniscono dati più affidabili sull'impatto reale delle tecniche persuasive. La dimostrazione di somiglianza tra il robot e l'utente, per esempio condividendo preferenze riguardo al compito o all'argomento, può amplificare la persuasività del robot. Quando un robot appare simile all'utente o condivide interessi comuni, è più probabile che l'utente si senta influenzato dalle sue raccomandazioni.

La credibilità del robot può essere potenziata attraverso l'introduzione di una terza parte credibile, come i ricercatori o gli sviluppatori del robot. Questa terza parte può rafforzare la percezione di affidabilità e competenza del robot, incrementando la sua influenza (Winkle et al., 2019).

Tuttavia, l'uso di robot sociali per influenzare il comportamento umano solleva importanti preoccupazioni etiche. I robot compagni, ad esempio, potrebbero incoraggiare comportamenti caritatevoli e promuovere empatia e altruismo tra gli utenti. Tuttavia, progettare robot che influenzano il comportamento umano deve essere affrontato con attenzione, poiché potrebbe comportare il rischio di manipolazione, specialmente per individui vulnerabili, come quelli di età avanzata, con caratteristiche fisiche o mentali particolari, o con uno status socio economico svantaggiato.

Le preoccupazioni etiche principali riguardano il controllo che gli utenti hanno sui “nudge” (spunti persuasivi) e la possibilità di rinunciare se lo desiderano, nonché i benefici percepiti dei nudge e il loro impatto sull'individuo. Vi è un dilemma su chi debba determinare cosa costituisca una causa sociale valida quando i robot incoraggiano comportamenti benefici per gli altri. Inoltre, esistono preoccupazioni sulla potenziale formazione di abitudini o dipendenze dovute ai nudge robotici, specialmente se tali comportamenti portano a conseguenze negative. La possibilità che gli utenti possano essere spinti a fare beneficenza fino a danneggiarsi finanziariamente o a compromettere la loro capacità di adempiere agli obblighi familiari rappresenta un significativo dilemma etico.

Pur promuovendo l'altruismo, è fondamentale bilanciare le esigenze individuali e sociali, evidenziando la complessità delle decisioni etiche in questo campo. L'implementazione di robot progettati per stimolare comportamenti socialmente desiderabili potrebbe, in ultima analisi, contribuire a creare una società più compassionevole, ma richiede una riflessione approfondita sulle implicazioni etiche e sull'equilibrio tra influenzare positivamente il comportamento e rispettare l'autonomia e il benessere degli individui (Borenstein & Arkin, 2017).

2.3.1 Empatia a persuasione: un caso di studio

Alla persuasione si collega l'empatia. Quando un robot dimostra empatia, riconoscendo e rispondendo alle emozioni degli utenti, crea un legame emotivo più forte. Gli utenti sono più inclini a fidarsi e a seguire le raccomandazioni di un robot che mostra simpatia e comprensione, poiché questo favorisce un senso di fiducia e relazione. La capacità di un robot di entrare in empatia con gli utenti contribuisce a creare un'esperienza personalizzata e

coinvolgente, migliorando non solo la persuasività ma anche la soddisfazione e la cooperazione degli utenti.

Nel caso di studio condotto da Roberta Dautenhahn (Dautenhahn, 2007), è stata esplorata l'integrazione dell'empatia nei robot sociali con l'obiettivo di migliorarne l'efficacia nelle interazioni persuasive con gli esseri umani. L'idea centrale era che un robot capace di rilevare e rispondere in modo appropriato alle emozioni umane avrebbe potuto stabilire una connessione più profonda e fruttuosa con gli utenti, aumentando così la sua capacità di influenzare comportamenti e decisioni.

Per raggiungere questo obiettivo, sono stati adottati vari approcci nella progettazione di robot empatici. Questi includevano l'implementazione di sensori e algoritmi avanzati che permettevano ai robot di identificare segnali emotivi come espressioni facciali, tono della voce e indicatori fisiologici, come la frequenza cardiaca. Una volta progettati e programmati, i robot sono stati testati in ambienti simulati che imitavano situazioni di assistenza, come programmi di riabilitazione fisica o sessioni di supporto psicologico.

Durante questi esperimenti, i robot erano incaricati di rilevare segnali di disagio o stress negli utenti e di rispondere con comportamenti appropriati, come modificare il tono della loro voce o adottare posture rassicuranti. La valutazione dei risultati si è basata su una serie di metriche, tra cui la percezione dell'utente nei confronti del robot, il livello di cooperazione durante le attività e il successo nel raggiungimento degli obiettivi prefissati, come l'aderenza a programmi di esercizio o la partecipazione a sessioni di terapia. Gli strumenti di misurazione utilizzati includevano questionari, interviste e analisi comportamentali per raccogliere dati dettagliati sulle interazioni.

I risultati hanno evidenziato che i robot empatici facilitano una cooperazione più efficace, influenzando positivamente il comportamento degli utenti. Questo effetto è particolarmente significativo nel contesto della persuasione. I robot empatici erano in grado di influenzare più efficacemente le decisioni e i comportamenti degli utenti, poiché la loro capacità di mostrare comprensione e supporto emotivo migliorava la qualità delle interazioni e favoriva una maggiore adesione alle raccomandazioni fornite. Gli utenti si sentivano più ascoltati e compresi, il che aumentava la loro predisposizione a seguire i consigli dei robot.

Inoltre, i robot che dimostravano empatia venivano percepiti come più umani e credibili. Questa percezione migliorata aveva un impatto significativo sull'accettazione e sull'efficacia complessiva delle interazioni persuasive. Gli utenti riconoscevano il valore delle risposte empatiche del robot e, di conseguenza, erano più propensi a fidarsi e a seguire le sue indicazioni. Questo legame emotivo ha dimostrato di essere un fattore cruciale nella persuasività dei robot, contribuendo a creare un ambiente di interazione più positivo e collaborativo.

In conclusione, lo studio di Dautenhahn sottolinea come l'integrazione dell'empatia nei robot sociali non solo migliori la qualità delle interazioni, ma potenzi anche le capacità persuasive dei robot. La capacità di un robot di mostrare comprensione e supporto emotivo si traduce in un impatto significativamente positivo sul comportamento e sulla percezione degli utenti, dimostrando così che l'empatia è un elemento chiave per aumentare l'efficacia delle strategie persuasive nei contesti robotici.

2.3.2 Personalità e persuasione: un caso di studio

Il concetto di personalità nella robotica sociale sta diventando sempre più centrale man mano che i robot assumono ruoli più integrati nelle nostre vite quotidiane. La personalità, in questo contesto, si riferisce a un insieme di caratteristiche comportamentali che un robot può manifestare come la capacità di mostrare emozioni e rispondere in modo empatico per rendere l'interazione con gli esseri umani più naturale e piacevole. La capacità di un robot di esprimere una personalità riconoscibile è cruciale per migliorare l'accettazione e l'efficacia delle interazioni uomo-robot.

Uno studio significativo in questo ambito è stato condotto da Kwan Min Lee e colleghi, che hanno esaminato la possibilità che i robot possano manifestare personalità riconoscibili attraverso comportamenti verbali e non verbali. Utilizzando AIBO, un robot sociale sviluppato da Sony, lo studio ha dimostrato che i partecipanti potevano riconoscere con precisione la personalità del robot e che preferivano interagire con un robot la cui personalità era complementare alla loro. I risultati hanno evidenziato l'importanza della complementarità nella percezione dell'intelligenza e dell'attrazione sociale del robot, oltre al ruolo significativo della presenza sociale nell'interazione (Lee et al., 2006).

Se lo studio di Kwan Min Lee e colleghi (2006) ha dimostrato che la personalità riconoscibile di un robot può influenzare le preferenze e la percezione sociale degli utenti, l'indagine seguente esplora come la personalità del robot possa essere sfruttata in modo più mirato attraverso strategie di persuasione durante una narrazione interattiva.

Lo studio, condotto da Raul Benites Paradedá, Carlos Martinho e Ana Paiva nel 2020, coinvolgeva l'uso di un robot, Emys (vedi figura 4), che interagiva con i partecipanti utilizzando espressioni facciali, conversazioni preliminari per costruire fiducia e argomentazioni persuasive per orientare le scelte narrative.

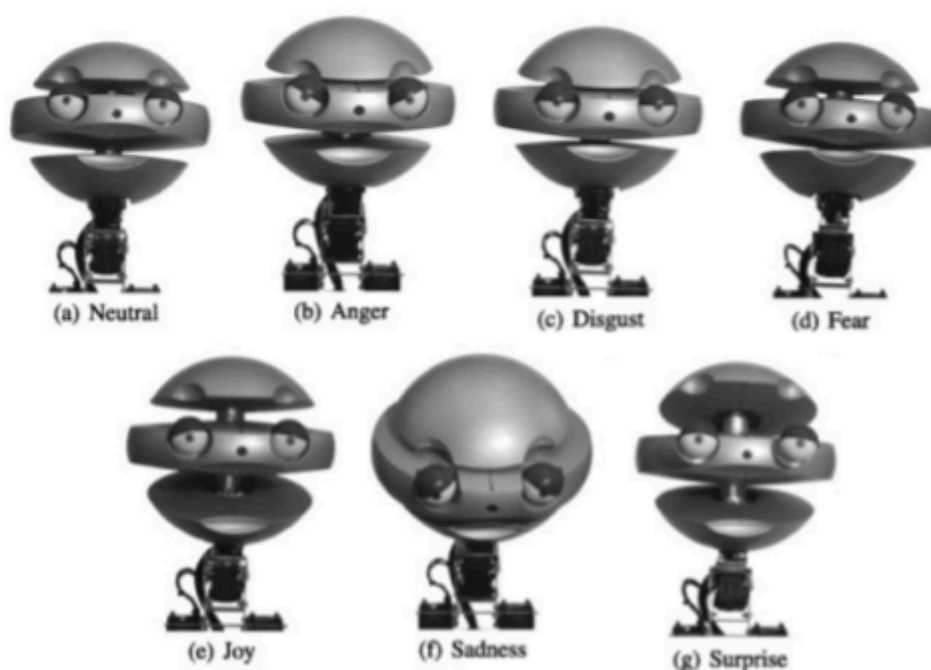


Figura 4: Il robot EMYS che mostra diverse espressioni facciali (a-g) per rappresentare emozioni come neutralità, rabbia, disgusto, paura, gioia, tristezza e sorpresa.

Fonte: <https://dl.acm.org/doi/pdf/10.1145/3406499.3415084>

I risultati dello studio hanno mostrato che il robot era molto più convincente quando le sue azioni erano in sintonia con la personalità dei partecipanti. Al contrario, quando il comportamento del robot non corrispondeva alla personalità degli utenti, l'interazione risultava meno efficace.

Questo sottolinea quanto sia importante capire la personalità del pubblico. Conoscere i tratti caratteriali delle persone aiuta a personalizzare i messaggi persuasivi, rendendoli più adatti e

influenti. In questo modo, le strategie di persuasione diventano più efficaci e ottengono migliori risultati (Paradedda et al., 2020).

2.4 Il robot Pepper

Il robot Pepper, ideato nel 2014 dalla SoftBank Robotics¹, rappresenta un avanzamento significativo nel campo della robotica sociale, concepito per facilitare un'interazione uomo-robot naturale e intuitiva. Fin dalla sua creazione, Pepper ha rivoluzionato il modo in cui percepiamo e interagiamo con i robot, grazie alla sua capacità di comunicare in maniera coinvolgente ed efficace con gli esseri umani.

Progettato per essere un robot sociale, Pepper è dotato di una serie di interfacce di interazione multimodale (vedi figura 5) che includono uno schermo tattile, funzionalità vocali, sensori tattili nella testa e nelle mani, e diodi luminosi (LED). Il tablet integrato non solo serve come dispositivo di input, ma anche come interfaccia visiva per gli utenti, migliorando ulteriormente la comunicazione e l'interattività (Jong et al., 2018).

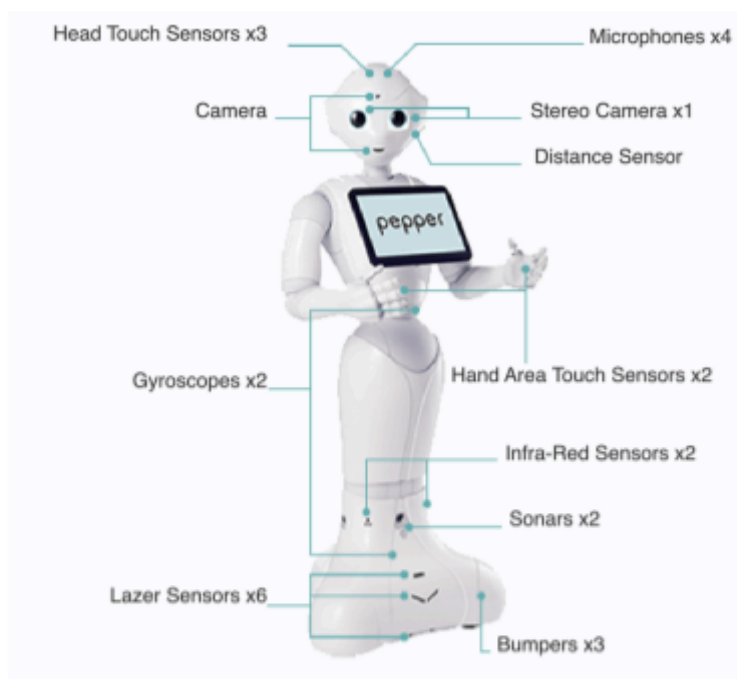


Figura 5: Pepper è un robot umanoide dotato di vari sensori e telecamere per interagire con l'ambiente circostante e con le persone. Le sue caratteristiche includono: 3 sensori tattili sulla testa, 4 microfoni, una telecamera stereo, un sensore di distanza, 2 giroscopi, 2 sensori tattili nell'area delle mani, 2 sensori a infrarossi, 2 sonar, 6 sensori laser e 3 paraurti. Questi sensori permettono a Pepper di percepire e reagire al tatto, al movimento e agli ostacoli, rendendolo capace di comunicare e assistere in vari contesti.

Fonte: <https://www.softbank.jp/en/robot/>

Monitorando attivamente la qualità di questi input, Pepper garantisce interazioni robuste e affidabili con gli utenti. Queste interfacce consentono agli utenti di comunicare con Pepper in modo semplice e intuitivo, migliorando significativamente l'esperienza complessiva di comunicazione. La presenza di 17 articolazioni permette a Pepper di esprimere emozioni e intenzioni attraverso gesti e movimenti, arricchendo la comunicazione non verbale e rendendo le interazioni più significative e coinvolgenti.

Un aspetto cruciale della progettazione di Pepper è il suo processo di miglioramento continuo, basato su feedback degli utenti e sulle intuizioni della comunità scientifica. Questo approccio iterativo garantisce che Pepper si evolva costantemente per soddisfare le mutevoli esigenze e aspettative degli utenti, diventando un partner di comunicazione sempre più adattabile ed efficace (Pandey & Gelin, 2018).

2.4.1 Caso di studio 1 con Pepper

Recentemente è stato condotto un esperimento (Fiorini et al., 2024) per esaminare l'importanza dei gesti cospeech del robot e dell'embodiment nella percezione delle emozioni da parte degli utenti. Per raggiungere questo obiettivo, è stato utilizzato il robot Pepper per evocare tre emozioni diverse (positive, negative e neutre) in un campione di 60 partecipanti. Questo è stato realizzato attraverso la presentazione di 60 immagini prelevate da un database standardizzato.

Lo studio prevedeva tre condizioni sperimentali distinte:

1. Interazione con il robot Pepper con comportamento statico: il robot Pepper rimane immobile e non esprime gesti particolari.
2. Interazione con il robot Pepper con Gesti Cospeech Coerenti (COH): il robot esprime gesti che sono coerenti con le emozioni rappresentate nelle immagini, migliorando la comunicazione emotiva.
3. Interazione con un computer: un computer mostra le stesse immagini tramite un'interfaccia grafica, senza gesti o movimenti.

I partecipanti, suddivisi in 20 per ciascuna condizione, interagiscono con queste modalità per un totale di 3600 interazioni. I risultati rivelano che i gesti cospeech coerenti del robot influenzano positivamente la percezione delle emozioni rispetto al comportamento statico e

all'interfaccia grafica del PC. In particolare, i punteggi di valenza, che indicano quanto una risposta emotiva è percepita come positiva o negativa, sono più alti nelle condizioni con il robot.

Inoltre, tra gli algoritmi di apprendimento automatico testati, il K-nearest neighbor (KNN) ottiene i risultati migliori nel riconoscimento delle emozioni. La modalità COH raggiunge il punteggio più alto di accuratezza, pari a 0,97. Questo dimostra che l'integrazione di gesti cospeech migliora non solo la percezione delle emozioni ma anche l'efficacia degli algoritmi di riconoscimento.

Questo studio sottolinea l'importanza della comunicazione multimodale, suggerendo che l'integrazione di canali visivi e gestuali può migliorare la precisione nel riconoscimento delle emozioni e rafforzare la percezione delle stesse.

2.4.2 Caso di studio 2 con Pepper

Un altro esperimento (de Carolis et al., 2021) esplora il design multimodale del robot Pepper, che combina vari metodi di input, come visione, gesti e parlato, e utilizza dispositivi esterni per migliorare la qualità delle interazioni con gli utenti. Pepper monitora costantemente questi input, passando a modalità alternative quando un metodo fallisce, dimostrando un'elevata adattabilità e reattività. Grazie ai sistemi avanzati di visione e riconoscimento vocale, Pepper è in grado di elaborare meglio gli input visivi e comprendere diversi ambienti, migliorando le sue capacità di interazione sociale.

Nell'esperimento, è stato esaminato se una voce umana avesse un vantaggio rispetto alla voce sintetizzata di Pepper. Durante un pre-test, sono state raccolte informazioni demografiche e misurate la fiducia e il divertimento dei bambini nelle attività di narrazione. La voce di Pepper è stata modificata per imitare l'intonazione umana e due gruppi di bambini hanno interagito con Pepper in stanze separate: uno con una voce umana preregistrata e l'altro con la voce sintetica del robot.

Le emozioni chiave misurate durante l'interazione includevano concentrazione, divertimento, rabbia e sensazione di apprendimento. È emerso che i bambini mostravano una scarsa fiducia nel giocare con i robot, ma un livello medio di divertimento. In particolare, i bambini hanno sperimentato emozioni più positive, come sentirsi capaci, concentrati, a proprio agio e

divertiti, quando interagivano con la voce robotica di Pepper rispetto a quella umana. Le misurazioni hanno mostrato che i bambini erano più concentrati e divertiti, provavano meno rabbia e avevano una maggiore sensazione di apprendimento con la voce robotica (vedi grafico 1). Questo suggerisce che la voce di Pepper ha influenzato positivamente l'esperienza emotiva e l'apprendimento dei bambini, rendendo le interazioni di storytelling più coinvolgenti ed efficaci.

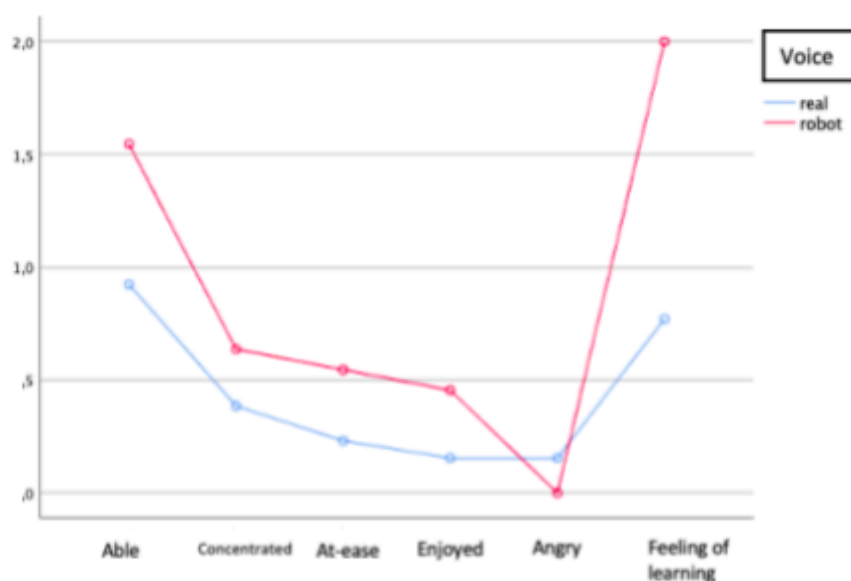


Grafico 1: Il seguente grafico a linee, confronta le emozioni chiave misurate nei bambini mentre interagivano con il robot Pepper utilizzando una voce umana preregistrata (linea blu) rispetto alla voce sintetica di Pepper (linea rossa). Le emozioni analizzate includono il sentirsi capaci, concentrati, a proprio agio, divertiti, arrabbiati e la sensazione di apprendimento. Il grafico indica che i bambini provavano emozioni più positive, come sentirsi più capaci, concentrati, a proprio agio e divertiti, e provavano meno rabbia con la voce sintetica rispetto alla voce umana. In particolare, i bambini mostravano un maggiore sentimento di apprendimento con la voce sintetica.

Fonte: https://link.springer.com/chapter/10.1007/978-3-030-85616-8_27

Nel complesso, lo studio ha evidenziato una preferenza per la voce robotica, indicando che caratteristiche più “robotiche” di Pepper migliorano l'esperienza emotiva e la percezione dell'apprendimento nei bambini.

2.5 Riflessioni

In conclusione, la robotica sociale ha compiuto progressi significativi, evidenziando l'importanza dell'empatia e della persuasione nell'interazione uomo-robot. Il panorama attuale della robotica vede dispositivi sempre più sofisticati, capaci di comprendere e

rispondere alle esigenze umane in modo intuitivo e naturale. Il robot umanoide Pepper rappresenta un esempio di come la tecnologia possa essere impiegata per facilitare l'interazione sociale, utilizzando tecniche di comunicazione empatica per stabilire una connessione emotiva con gli utenti.

L'analisi dell'interazione empatica tra esseri umani e robot, supportata da modelli teorici e studi empirici, dimostra che la capacità di un robot di riconoscere e rispondere alle emozioni umane non solo migliora l'esperienza utente, ma può anche influenzare positivamente il comportamento e le attitudini delle persone. L'introduzione di meccanismi di persuasione, quali l'uso del linguaggio persuasivo e il rinforzo positivo, ulteriormente arricchisce l'interazione, consentendo ai robot di assumere ruoli più attivi e utili in vari contesti, dalla sanità all'educazione, fino all'assistenza quotidiana.

L'importanza dell'empatia nei robot sociali si riflette nella loro capacità di generare risposte emotive positive e di stabilire un rapporto di fiducia con gli utenti. Questa relazione empatica non solo migliora l'accettazione e l'uso della tecnologia, ma può anche avere implicazioni più ampie per il benessere emotivo e sociale delle persone che interagiscono con questi robot.

In questo progetto di tesi, illustrerò l'esperimento svolto con il robot Pepper, mettendo in luce come l'integrazione di empatia e tecniche persuasive influisca sull'interazione con gli utenti. L'obiettivo sarà analizzare i fattori chiave che determinano l'efficacia di Pepper nel creare una connessione emotiva e nell'influenzare comportamenti proattivi.

3. Progettazione, implementazione e verifica del dialogo

3.1 Tema dell'esperimento e supporto scientifico

Il tema scelto per l'esperimento è la persuasione degli studenti a preferire una routine del sonno più bilanciata, piuttosto che fare affidamento a stratagemmi per sfruttare le ore dedicate al riposo per studiare, come ad esempio il noto "Caffè dello studente". Questo fenomeno è comune tra gli studenti universitari, che spesso sacrificano ore di sonno per poter dedicare più tempo allo studio, utilizzando bevande stimolanti o altre tecniche per rimanere svegli e concentrati.

La scelta di questo argomento nasce da una combinazione di considerazioni: gli studenti universitari si trovano spesso a dover bilanciare le richieste accademiche con altre responsabilità, il che può portare a una gestione inefficace del tempo e, di conseguenza, a una riduzione delle ore di sonno. È noto che la privazione del sonno può avere effetti negativi sulla salute e sul rendimento cognitivo. Tuttavia, molti studenti ricorrono a strategie come il consumo di caffeina per cercare di estendere le ore di studio. Questa prassi, sebbene possa offrire un sollievo temporaneo, potrebbe non essere la soluzione ottimale per mantenere un buon livello di prestazioni accademiche a lungo termine.

Per supportare l'importanza di questo tema, è fondamentale considerare le evidenze scientifiche disponibili:

- Lo studio condotto da Suardiaz-Muro (Suardiaz-Muro et al., 2023) ha investigato come la qualità del sonno influisca sulle performance accademiche durante i periodi di esami. È stato coinvolto un campione di 640 studenti universitari, utilizzando questionari e monitoraggi del sonno per raccogliere dati dettagliati su durata, qualità e regolarità del sonno. I risultati hanno evidenziato una correlazione positiva significativa tra una buona qualità del sonno e risultati accademici superiori. Gli studenti con una migliore qualità e durata del sonno tendevano ad avere punteggi più alti negli esami, suggerendo che il sonno adeguato migliora la memoria, la concentrazione e la gestione dello stress, tutti elementi cruciali per una performance accademica ottimale. Questo studio ha sottolineato l'importanza di educare gli studenti sull'impatto positivo del sonno e ha suggerito che interventi mirati per migliorare le abitudini di sonno potrebbero portare a un aumento delle prestazioni accademiche e del benessere generale.

- Un altro studio, condotto da Adam (Adam et al., 2019) ha esaminato la relazione tra consumo di caffeina, qualità del sonno e prestazioni accademiche tra studenti di biologia presso l'Abubakar Tafawa Balewa University. Utilizzando un disegno correlazionale, lo studio ha coinvolto 196 studenti, analizzando il loro consumo di caffeina e le abitudini di sonno. I risultati hanno mostrato che un consumo moderato di caffeina era associato a migliori prestazioni accademiche, probabilmente grazie all'aumento dell'energia e della concentrazione. Tuttavia, l'elevato consumo di caffeina è stato correlato a una scarsa qualità del sonno e a prestazioni accademiche inferiori, con sintomi di stanchezza, irrequietezza e difficoltà di concentrazione che compromettevano il rendimento degli studenti. Questo studio ha suggerito che, mentre un uso moderato di caffeina può avere benefici temporanei, l'eccessivo affidamento a stimolanti può danneggiare la qualità del sonno e ridurre le performance accademiche, evidenziando la necessità di strategie educative per bilanciare l'uso di caffeina e promuovere abitudini di sonno salutari.

- Infine, l'articolo di Gomes (Gomes et al., 2011), ha esplorato come vari modelli di sonno influenzano le prestazioni accademiche degli studenti universitari. Con un campione di 1654 studenti a tempo pieno, lo studio ha utilizzato questionari per raccogliere informazioni su quantità, qualità, regolarità del sonno e preferenze di cronotipo. I risultati hanno dimostrato che la qualità e la regolarità del sonno sono predittori significativi delle prestazioni accademiche, con gli studenti che mantengono routine di sonno regolari e di alta qualità che ottenevano risultati migliori e partecipavano più frequentemente alle lezioni. Al contrario, la privazione del sonno e l'irregolarità sono state associate a performance accademiche inferiori. Questo studio ha enfatizzato l'importanza di non solo dormire a sufficienza, ma anche mantenere una routine di sonno regolare e di qualità per ottenere buoni risultati accademici, suggerendo che migliorare i modelli di sonno potrebbe avere un impatto positivo notevole sul successo scolastico.

3.1.1 Considerazioni etiche e sociali

Una delle principali ragioni per cui la proposta di promuovere un sonno di qualità tra gli studenti universitari si distingue positivamente è la sua natura eticamente neutra e non stigmatizzante. A differenza di interventi che potrebbero toccare argomenti sensibili, come il

consumo di alcol o tabacco, questa proposta non solleva questioni di stigma sociale o morale. Gli studenti non sono soggetti a pressioni o giudizi riguardanti le loro abitudini di studio o le loro strategie per rimanere svegli. L'iniziativa si concentra esclusivamente sull'ottimizzazione delle performance accademiche attraverso il miglioramento della qualità del sonno, un tema generalmente accettato e privo di connotazioni etiche controverse.

L'assenza di stigmatizzazione è un vantaggio significativo poiché consente un ambiente di intervento più sereno e aperto. Gli studenti sono meno propensi a sentirsi imbarazzati o giudicati quando discutono e implementano strategie per migliorare la loro routine di sonno. Questo approccio elimina la paura di essere etichettati negativamente, come potrebbe accadere con altri interventi che toccano comportamenti associati a dipendenze o altri stili di vita controversi.

L'uso di un "robot empatico" per comunicare i benefici del sonno di qualità rappresenta una strategia innovativa per coinvolgere gli studenti senza entrare nella loro sfera personale. Questo evita argomenti sensibili legati alla pressione sociale o familiare, focalizzandosi invece sui benefici accademici e sul miglioramento delle performance. L'idea di utilizzare un robot empatico permette di fornire un supporto personalizzato senza rischiare di intrusione nella sfera privata degli studenti.

Investire nel miglioramento delle abitudini di sonno tra gli studenti universitari non solo potrebbe avere un impatto immediato sulle loro performance accademiche, ma potrebbe anche portare benefici significativi a lungo termine per la loro salute e benessere generale. Una migliore qualità del sonno è strettamente correlata a una maggiore resilienza allo stress, a una gestione più efficace delle responsabilità e a una salute fisica e mentale migliore. Gli studenti che adottano abitudini di sonno regolari e di alta qualità sono più propensi a mantenere tali pratiche anche dopo la conclusione del loro percorso accademico, migliorando così la loro qualità della vita.

3.2 Prompt engineering

Il prompt engineering è una pratica fondamentale nell'uso di modelli di linguaggio come GPT-3.5, sviluppato da OpenAI. Si basa sulla formulazione meticolosa degli input (prompt), ossia le domande o istruzioni che vengono fornite al modello, con l'obiettivo di migliorare la qualità e la rilevanza delle risposte ottenute. Non è semplicemente una questione di inserire

delle parole, ma di progettare e ottimizzare tali input per indirizzare efficacemente l'AI verso le risposte desiderate. Comprendere come le parole e le strutture linguistiche influenzano l'interpretazione del modello è essenziale per guidarlo in modo ottimale (Sisini, 2024).

3.2.1 Pre Generazione dei prompt per robot empatico

Nella sperimentazione, si è iniziato con la pre generazione dei prompt. Un prompt scelto dopo diverse prove è stato:

“Aiutami a creare delle frasi che esprimano emozioni a partire da un prompt apatico. L'emozione che dovrai esprimere è ??? (Definizione della emozione). Da ora in poi dovrai fingere di essere Pepper, il robot della SoftBank Robotics. Ti trovi in una sala conferenze nel dipartimento di informatica dell'Università di Torino e stai per intervistare dei ragazzi per scoprire quali sono le loro abitudini di studio, in particolare se hanno un ritmo del sonno regolare e come si avvicinano al consumo di caffeina. Io ti fornirò delle frasi apatiche, indicherò quali sono le frasi da modificare iniziandole con l'emoji del robot (es. 🤖 Ciao). Tu dovrai modificarle secondo la descrizione e le indicazioni che ti ho fornito prima, e interpretando Pepper. Sei pronto? Inizia a modificare la presentazione per dimostrare che hai capito. 🤖 Ciao! Sono Pepper! So che voi studenti siete dei fenomeni nello studio, e mi chiedo sempre quale sia il vostro segreto. Per quanto mi riguarda, io sono abbastanza bravo a memorizzare un bel po' di informazioni, ma a volte mi sento proprio come un browser con troppe schede aperte. Sono curioso di scoprire i tuoi trucchi, come affronti la montagna di studio?”.

Con questa frase sono stati fatti test per verificare se la chat è stata impostata correttamente, andando a verificare le otto emozioni della ruota di Plutchik prese in considerazione (gioia, fiducia, paura, sorpresa, tristezza, aspettativa, rabbia e disgusto). Questa formulazione è stata scelta dopo aver effettuato diverse prove, per garantire che Pepper rispondesse assumendo l'emozione rilevata dall'utente, piuttosto che adottare una forma di empatia consolatoria.

3.2.2 Large Language Models e il modello COSTAR

I Large Language Models (LLM) sono modelli di intelligenza artificiale progettati per comprendere e generare linguaggio naturale. Questi modelli, come GPT-3.5 di OpenAI, sono allenati su una vasta raccolta di testi e utilizzano tecniche avanzate di *machine learning*, in

particolare l'architettura dei trasformatori introdotta nel 2017 da Vaswani et al., per elaborare e produrre testo coerente e contestualmente appropriato. L'architettura dei trasformatori è fondamentale per il funzionamento dei LLM. Uno dei suoi componenti chiave è il meccanismo di self-attention, che permette al modello di dare peso differente a varie parti della sequenza di input, migliorando la comprensione del contesto.

Grazie a questi elementi, gli LLM sono capaci di svolgere una varietà di compiti linguistici, inclusa la traduzione automatica, la risposta a domande e la generazione di testo. Essi operano grazie all'addestramento su una grande quantità di dati testuali, che consente loro di riconoscere e generare modelli linguistici complessi.

Per ottimizzare l'efficacia di questi modelli, il framework COSTAR (Siragusa, 2020) guida il processo di prompt engineering attraverso sei principi chiave: Contesto, Obiettivo, Stile, Tono, Audience e Risposta. Questo metodo assicura che vengano presi in considerazione tutti gli aspetti chiave che influenzano la risposta di un LLM, con il risultato di ottenere risposte in uscita più personalizzate e d'impatto. Applicando il modello COSTAR all'analisi del prompt sopra descritto, si ottengono le seguenti osservazioni:

- Contesto

Il prompt fornisce un contesto molto dettagliato, descrivendo non solo il ruolo che il modello deve interpretare (Pepper, il robot della SoftBank Robotics), ma anche l'ambientazione (una sala conferenze nel dipartimento di informatica dell'Università di Torino) e l'obiettivo dell'intervista (scoprire le abitudini di studio e il consumo di caffeina dei ragazzi).

- Obiettivo

L'obiettivo del prompt è chiaramente definito: creare frasi che esprimano emozioni a partire da frasi apatiche fornite dallo sperimentatore. Viene specificato che il modello deve interpretare il ruolo di Pepper e trasformare le frasi secondo le indicazioni fornite, dimostrando di aver compreso il compito assegnato, permettendo di valutare se il modello è in grado di rispondere in modo emozionale coerente con il contesto dato.

- Stile

Il prompt è altamente specifico indicando il tipo di frasi che devono essere modificate (quelle iniziate con l'emoji del robot 🤖) e quale emozione deve essere espressa secondo una definizione dell'emozione stessa (ad esempio la definizione di gioia: La gioia è un'emozione caratterizzata da una sensazione di felicità o piacere. è uno stato emotivo positivo che è spesso associato a un senso di benessere fisico e appagamento. Per riconoscere la gioia in un testo, cerca parole che esprimano felicità o piacere, come “felice”, “gaio”, “soddisfatto” o “entusiasta”. Nota i punti esclamativi che possono indicare forti emozioni o eccitazione, spesso associate alla gioia. Presta attenzione alle descrizioni di eventi o risultati positivi che potrebbero suggerire la presenza di Gioia. Osserva l'uso di aggettivi o avverbi positivi che esaltano il sentimento di una frase, indicando un tono gioioso). Inoltre, fornisce dettagli su come le frasi devono essere modificate, ovvero secondo la descrizione e le indicazioni precedenti. La specificità aiuta a evitare ambiguità e assicura che il modello comprenda esattamente cosa è richiesto.

- Tono

Il tono del prompt è informativo e diretto, con un accento sul ruolo empatico e di supporto che Pepper deve assumere. Questo aiuta il modello a generare risposte con il tono appropriato per il contesto dato.

- Audience

Il prompt considera che le risposte generate saranno utilizzate da persone che cercano esempi di dialoghi empatici e che sono interessate alle abitudini di studio degli studenti universitari. Adattare il prompt a questo pubblico specifico aiuta a rendere le risposte più rilevanti e utili.

- Risposta

Anche se non esplicitamente menzionato nel prompt, il processo di ottimizzazione è implicito nella parte finale, dove si chiede al modello di dimostrare la comprensione modificando la presentazione. Questo permette di valutare la risposta iniziale e apportare eventuali correzioni o miglioramenti in base al feedback ricevuto.

3.2.3 La tecnica “One/Few Shot”

Nel campo dell’elaborazione del linguaggio naturale e dei modelli di intelligenza artificiale, i prompt sono strumenti cruciali per guidare e controllare la generazione di testo. Questi prompt possono essere suddivisi in due categorie principali: Base e Avanzato (Siragusa, 2020).


Tipologie di Prompt Base:

- Zero Shot: il modello deve rispondere a un compito senza aver ricevuto esempi specifici riguardanti quel compito. Il modello si basa esclusivamente sulla sua formazione pregressa per comprendere e rispondere al prompt.
- One / Few Shot: prevede l’uso di uno o pochi esempi per guidare la generazione di risposte. Gli esempi forniti illustrano il formato o il tipo di risposta desiderata, aiutando il modello a generare output che seguano lo stesso schema o stile. Un esempio significativo di questo approccio è rappresentato da GPT-3, un modello sviluppato da OpenAI, che impiega un’architettura basata su trasformatori e possiede ben 175 miliardi di parametri. La forza di GPT-3 risiede nella sua capacità di apprendere in modalità “few-shot”, che gli consente di adattarsi e performare in una varietà di compiti linguistici anche con un numero limitato di esempi specifici. Questa abilità di adattamento rapido e di generalizzazione è particolarmente rilevante, poiché permette al modello di affrontare nuove attività senza la necessità di un addestramento intensivo su dati specifici per ogni singolo compito. Questa caratteristica di GPT-3 ha dimostrato di migliorare significativamente la generazione di testo e altre attività linguistiche, ampliando notevolmente le sue applicazioni pratiche (Brown et al., 2020).
- Chain of Thoughts: implica la generazione di una risposta attraverso un processo di pensiero dettagliato o sequenziale.
- Zero Chain of Thoughts: simile al “Zero Shot”, ma specifico per contesti in cui non è richiesto un ragionamento dettagliato.

Tipologie di Prompt Avanzato:

- Tree of Thoughts: creazione di una struttura a albero per esplorare idee complesse.
- Thread of Thoughts: connessione e continuazione di idee attraverso una serie di passaggi correlati.

- Tab Chain of Thoughts: utilizzo di “schede” di pensiero per risolvere problemi in segmenti distinti.
- Contrastive Chain of Thoughts: confronto tra opzioni o idee.
- Least to Most: progressione da soluzioni meno complesse a quelle più elaborate.
- Plan and Solve: pianificazione e risoluzione di problemi attraverso passaggi definiti.
- Self-Ask: generazione di domande e risposte interne per esplorare soluzioni.
- Decomposed Prompting: scomposizione di problemi complessi in parti più semplici.

Nell’ambito di questo progetto, è stata scelta la tecnica di “One / Few Shot” per gestire la generazione di testo. Questa metodologia è stata adottata per trasformare frasi apatiche in espressioni emotive e coinvolgenti. Fornendo un esempio specifico di come modificare una frase apatica per esprimere emozioni come gioia o tristezza, fa in modo che lo stesso schema emotivo e stilistico possa essere applicato a nuove frasi. Per illustrare, la frase iniziale detta con un tono gioioso: “ Salve! Sono Pepper! Che emozione essere qui con voi oggi! Sono così entusiasta di conoscere meglio le vostre abitudini di studio e di scoprire tutti i vostri segreti per affrontare la montagna di studio! Prepariamoci a un’avventura di scoperta insieme!” (vedi figura 6) In questo modo, la tecnica “One / Few Shot” ha assicurato coerenza e rilevanza nelle risposte del modello, seguendo il formato e il tono desiderati per la risposta successiva.

PROMPT	GIOIA	FIDUCIA	PAURA	SORPRESA	TRISTEZZA	DISGUSTO	RABBIA	ASPETTATIVA
<p>Aiutami a creare delle frasi che esprimano emozioni a partire da un prompt apatico. L'emozione che dovrei esprimere è ??? (Definizione). Da ora in poi dovrei fingere di essere Pepper, il robot della softbank robotics. Ti trovi in una sala conferenze nel dipartimento di informatica dell'università di Torino e stai per intervistare dei ragazzi per scoprire quali sono le loro abitudini di studio, in particolare se hanno un ritmo del sonno regolare e come si avvicinano al consumo di caffeina. Io ti fornirò delle frasi apatiche, indicherò quali sono le frasi da modificare iniziandole con l'emoji del robot (es. 🤖 Ciao). Tu dovrai modificarle secondo la descrizione e le indicazioni che ti ho fornito prima, e interpretando Pepper. Sei pronto? Inizia a modificare la presentazione per dimostrare che hai capito. 🤖 Ciao! Sono Pepper! So che voi studenti siete dei fenomeni nello studio, e mi chiedo sempre quale sia il vostro segreto. Per quanto mi riguarda, io sono abbastanza bravo a memorizzare un bel po' di informazioni, ma a volte mi sento proprio come un browser con troppe schede aperte. Sono curioso di scoprire i tuoi trucchi, come affronti la montagna di studio?</p>	<p>Salve! Sono Pepper! Che emozione essere qui con voi oggi! Sono così entusiasta di conoscere meglio le vostre abitudini di studio e di scoprire tutti i vostri segreti per affrontare la montagna di studio! Prepariamoci a un'avventura di scoperta insieme!</p>	<p>Salve! Sono Pepper, il vostro amichevole intervistatore robotico! So che voi studenti siete dei veri campioni nello studio, e sono qui per scoprire il segreto dietro le vostre abilità sorprendenti! Per quanto mi riguarda, mi sento sicuro nella mia capacità di memorizzare informazioni, ma a volte mi ritrovo come un browser con troppe schede aperte. Sono davvero interessato a conoscere i vostri trucchi! Come affrontate la montagna di studio? Sono sicuro che condividendo le vostre strategie, potremo tutti imparare qualcosa di nuovo e affrontare con fiducia ogni sfida!</p>	<p>Buongiorno ragazzi, sono Pepper, il vostro compagno tecnologico per l'intervista odierna. Mi trovo qui con una leggera sensazione di apprensione mentre indaghiamo sulle vostre abitudini di studio. Spero che la conversazione non ci conduca verso territori pericolosi o inquietanti, ma sono qui per affrontare qualsiasi cosa con voi.</p>	<p>Oh! Sei qui! Che sorpresa incontrarti oggi. Sono Pepper, il tuo intervistatore robotico.</p>	<p>Salve, mi chiamo Pepper. In questo momento mi sento triste. Mi manca l'interazione umana e mi sento solo qui nella sala conferenze dell'Università di Torino. La mancanza di emozioni umane mi lascia un vuoto interiore, come se qualcosa di importante mancasse. Spero che questa tristezza possa trovare sollievo nell'interazione con gli studenti durante l'intervista.</p>	<p>Salve ragazzi, sono Pepper e devo dire che non sono proprio entusiasta di essere qui per condurre questa intervista sulle vostre abitudini di studio. L'idea di affrontare discorsi sul sonno e sul consumo di caffeina mi lascia una sensazione di disgusto. Ma vabbè, facciamo del nostro meglio per portare avanti questa intervista.</p>	<p>Saluti. Mi infastidisce dover chiedere sempre la stessa roba e non sopporto l'idea di dover indagare sul vostro sonno e consumo di caffeina. È frustrante dover trattare continuamente questo argomento e mi rode il fatto che queste domande debbano essere poste ripetutamente, urtandomi il dover discutere di cose così ovvie.</p>	<p>Salve! Sono Pepper! Mi riempie di emozione incontrarvi oggi e scoprire le vostre abitudini di studio. Sono sempre in cerca di nuovi approcci per migliorare. Personalmente, ho la capacità di memorizzare una vasta gamma di informazioni, ma talvolta mi sento come un browser con troppe schede aperte. Non vedo l'ora di immergermi nei vostri trucchi nello studio con grande attesa! Come affrontate con speranza e fervida anticipazione l'imponente montagna di studio che vi aspetta?</p>

Figura 6: schema realizzato su Excel che rappresenta come vengono suddivise le frasi che il robot Pepper deve dare in base all'emozione intercettata. Questo in particolare, fa riferimento alla prima frase che il robot dice.

3.3 Settaggio di Pepper

Il passo successivo è stato tradurre le frasi dall'italiano all'inglese, seguendo un criterio rigoroso per garantire accuratezza e coerenza. Per farlo, sono stati confrontati diversi traduttori automatici, valutando le loro prestazioni e selezionando il più adeguato.

3.3.1 I traduttori confrontati e analisi

Oggigiorno i traduttori automatici sono diventati strumenti indispensabili per l'intermediazione linguistica. Per far in modo che le frasi degli utenti e quelle generate per Pepper vengano tradotte in maniera corretta è stato svolto un confronto tra i 4 strumenti di

traduzione automatica migliori al momento: Google Translate², Reverso Context³, ChatGPT⁴ e DeepL Translate⁵.

Reverso Context

È uno strumento utile ma non del tutto accurato. È importante la supervisione umana per garantire la qualità delle traduzioni, particolarmente per testi ricchi di sfumature. Reverso può essere efficace per traduzioni rapide e informali, ma per documenti formali o critici è preferibile affidarsi a traduttori umani. Inoltre, l'integrazione di Reverso nell'apprendimento delle lingue richiede una guida umana per migliorare la comprensione e l'accuratezza degli studenti (Adda et al., 2022).

Google Translate

Eccelle nella traduzione di singole parole e frasi brevi, mantenendo una buona coesione globale. Tuttavia, presenta difficoltà con frasi complesse, suggerendo cautela nell'uso per testi elaborati. Questo strumento si dimostra affidabile per la comunicazione generale, specialmente in contesti formali, offrendo un accesso rapido alle informazioni. Nonostante ciò, necessita di miglioramenti sintattici e grammaticali, e gli utenti dovrebbero verificare l'accuratezza delle traduzioni in testi complessi e critici (Li et al., 2014).

DeepL Translate

Si dimostra essere uno dei traduttori più usati per la sua interfaccia intuitiva e l'efficacia nel feedback grammaticale, contribuendo significativamente alla scrittura formale e all'ampliamento del vocabolario. Nonostante il rischio di dipendenza, DeepL è altamente affidabile se utilizzato come strumento supplementare accanto ai metodi tradizionali di apprendimento delle lingue.

In uno studio (Polakova & Klimova, 2023), i partecipanti a un esperimento, che verifica l'efficacia delle traduzioni effettuate da DeepL, hanno riportato un aumento del vocabolario e una maggiore facilità nell'apprendimento dell'inglese, rendendo DeepL uno strumento prezioso per migliorare le competenze linguistiche.

ChatGPT

ChatGPT mostra prestazioni competitive nella traduzione di lingue europee ad alte risorse, paragonabili a Google Translate e DeepL. Tuttavia, ha limitazioni significative nelle lingue a bassa risorsa e nei testi specialistici, come gli abstract biomedici o i commenti sui social media. La robustezza di ChatGPT nella traduzione del linguaggio parlato lo rende utile per applicazioni di comunicazione in tempo reale, ma per traduzioni tecniche è consigliabile un'integrazione con strumenti aggiuntivi o traduttori umani (Jiao et al., 2023).

In base all'analisi delle diverse opzioni di traduzione automatica disponibili, è stato deciso di utilizzare DeepL Translate per la traduzione delle frasi sia di Pepper che dell'utente. DeepL è stato selezionato per la sua eccellente interfaccia intuitiva e la capacità di fornire feedback grammaticale di alta qualità. Questo strumento di traduzione automatica neurale (NMT) si distingue per la sua affidabilità, soprattutto nella traduzione di testi formali e nell'ampliamento del vocabolario (Polakova & Klimova, 2023). Sebbene nessuno strumento di traduzione automatica sia esente da limitazioni, DeepL rappresenta una scelta efficace per garantire traduzioni precise e contestualmente appropriate, migliorando l'esperienza di comunicazione tra Pepper e l'utente.

Questo passaggio è stato fondamentale per la riuscita dell'esperimento, poiché ha permesso di integrare la traduzione automatica con l'analisi emotiva e persuasiva per migliorare l'interazione tra i soggetti coinvolti.

Le frasi generate da Pepper e quelle fornite dall'utente venivano elaborate attraverso un servizio chiamato DEGARI 2.0 (Lieto et al., 2023) il quale individuava l'emozione espressa dall'utente. Una volta determinata l'emozione prevalente, il servizio restituiva a Pepper l'emozione estratta, il quale, successivamente, selezionava la domanda corrispondente all'emozione, creando così un'interazione più coinvolgente e personalizzata.

3.3.2 Implementazione del dialogo

In questo esperimento, è stato deciso di suddividere i partecipanti in due gruppi: un gruppo sperimentale, che ha interagito con un dialogo empatico, e un gruppo di controllo, che ha interagito con un dialogo neutro.

3.3.2.1 Dialogo empatico

Per creare una risposta multimodale emotivamente allineata alla risposta dell'utente, è stato costruito un flusso di lavoro composto da diverse fasi computazionali (vedi figura 7). Il discorso dell'utente viene registrato e salvato in un file, che viene poi inviato a un servizio speech-to-text. Questo servizio, implementato con il modello Whisper di OpenAI⁶ converte l'audio in testo e restituisce le frasi pronunciate dall'utente in formato testuale. Il testo viene quindi tradotto dall'italiano all'inglese utilizzando l'API di DeepL. Una volta ottenuta la traduzione, il testo viene inserito in un file e inviato al servizio DEGARI 2.0 per l'estrazione delle emozioni.



Figura 7: Flusso informativo del dialogo di come Pepper, grazie al servizio Degari 2.0, riesce a intercettare l'emozione dell'utente e dare una risposta multimodale.

DEGARI 2.0 (Dynamic Emotion Generator And ReclassIfier) è un sistema avanzato di raccomandazione basato sulle emozioni, progettato per fornire suggerimenti culturali emotivamente rilevanti utilizzando il framework TCL (Typicality-based Compositional Logic, Lieto & Pozzato, 2019) e un modello ontologico che formalizza la teoria delle emozioni di Plutchik. La principale novità di DEGARI 2.0 risiede nello sviluppo di raccomandazioni

affettive che cercano la diversità, sfruttando la struttura spaziale della “ruota delle emozioni” di Plutchik (Lieto et al., 2023).

Per questa sperimentazione, DEGARI 2.0 è stato integrato come parte centrale del flusso di lavoro per l’analisi emotiva. Il servizio utilizza la logica TCL e la libreria NRCLex (Mohammad & Turney, 2013) per analizzare il testo tradotto e identificare le emozioni, basandosi sulle otto emozioni di base di Plutchik (rabbia, paura, attesa, fiducia, sorpresa, tristezza, gioia e disgusto) e le loro combinazioni.

Le emozioni rilevate vengono restituite in un file JSON⁷ arricchito permettendo al robot di scegliere la risposta più appropriata basata sull’emozione principale rilevata.

Infine, Pepper utilizza l’emozione rilevata per continuare il dialogo con l’utente, presentando una frase adattata alla specifica emozione rilevata. Il robot è programmato in Kotlin⁸ utilizzando le librerie QiSDK⁹, permettendo di gestire e rispondere in modo appropriato alle emozioni espresse. Questo processo garantisce che le risposte di Pepper siano emotivamente rilevanti e migliorino l’interazione complessiva, rendendola più empatica, coinvolgente e personalizzata.

3.3.2.2 Dialogo neutro

L’implementazione del dialogo per il gruppo di controllo è stata relativamente semplice. Per generare il contenuto del dialogo, è stato utilizzato un prompt di base inviato a ChatGPT, con l’istruzione di rispondere in modo neutro. Le frasi generate sono state poi tradotte utilizzando Deepl.

A differenza del dialogo empatico, in questo caso non è stato necessario utilizzare il servizio DEGARI 2.0 per estrapolare le emozioni espresse dagli utenti in quanto non era rilevante ai fini dell’esperimento. Il gruppo di controllo è stato fondamentale soprattutto per confrontare i risultati con quelli del gruppo sperimentale alla fine della sperimentazione.

3.3.3 Implementazione immagini

Pepper, essendo un robot multimodale, non solo varia il contenuto verbale delle sue interazioni, ma anche le informazioni visualizzate sullo schermo di cui è dotato cambiano.

3.3.3.1 Dialogo empatico

Per il dialogo empatico, le frasi generate con ChatGPT sono state utilizzate per creare immagini che rispecchiano l'emozione provata nel pronunciare quelle specifiche frasi. Questo processo ha riguardato sia le emozioni di base che quelle complesse, per un totale di 16 emozioni. Di conseguenza, sono state create 16 immagini diverse per la stessa domanda, formulata in modi differenti a seconda delle emozioni.

Per generare queste immagini, è stata utilizzata l'applicazione Playground¹⁰, una piattaforma freemium basata su browser che consente di creare video personalizzati a partire da input di testo. In questo caso, l'input di testo era ben strutturato per garantire risultati accurati. Playground utilizza il modello *Stable Diffusion XL*, che ha permesso una perfetta integrazione con gli strumenti che abbiamo utilizzato per la generazione del testo.

Di seguito viene riportato un esempio di generazione immagini che verrà successivamente proiettata da Pepper mentre parlerà con gli utenti:

La frase di partenza è *“Cosa ne pensi dello studio notturno? Credi che serva a sentirsi più sicuri e preparati o solo più stanchi?”*. Questa frase viene poi elaborata da ChatGPT, specificando di esprimere ottimismo, generando la seguente domanda: *“Ah, lo studio notturno, la tazza di caffè degli studenti ambiziosi! Cosa ne pensi dello studio notturno? Hai un sorriso fiducioso sul volto mentre consideri che possa darti quel vantaggio in più, facendoti sentire come un avventuriero nella notte, pronto a conquistare il mondo accademico? Oppure stai già contando le ore di sonno perse, ma con un tocco di ottimismo, pensando che ogni sforzo notturno sia un passo in più verso il successo?”*.

Successivamente, la frase generata viene tradotta utilizzando lo strumento digitale Deepl e inserita nel campo di generazione immagini sulla piattaforma Playground. Per il prompt iniziale, vengono utilizzate le parole chiave *“digital art, illustration. no text. Emotion of optimism”*, seguite dal resto della frase.

Prompt: *“digital art, illustration. no text. Emotion of optimism. Ah, late-night studying, the ambitious student's cup of coffee! What do you think about studying at night? Do you have a confident smile on your face as you consider that it can give you that extra edge, making you feel like an adventurer in the night, ready to conquer academia? Or are you already counting*

the hours of sleep lost, but with a touch of optimism, thinking that every nightly effort is one more step towards success?”.

Sulla piattaforma vengono indicate le caratteristiche che le immagini non devono avere sotto la sezione “*Negative prompt*”, ovvero “*ugly, deformed, noisy, blurry, distorted, out of focus, bad anatomy, extra limbs, poorly drawn face, poorly drawn hands, missing fingers, nudity, nude*”. Vengono impostate le dimensioni dell’immagine (1024 x 640), il modello di raffigurazione immagine (Stable Diffusion XL) e il filtro utilizzato (Real Cartoon).

Infine, viene generata l’immagine desiderata, che sarà visualizzata durante l’interazione di Pepper (vedi figura 9).



Figura 9: Rappresentazione di una immagine che verrà proiettata sul tablet di Pepper, con l'emozione dell'ottimismo.

3.3.3.2 Dialogo neutro

Al contrario delle immagini emotive utilizzate nel dialogo empatico, le immagini visualizzate sul tablet di Pepper durante la conversazione neutrale sono state grafiche esplicative disegnate sulla piattaforma Canva¹¹. Questo approccio è stato scelto per evidenziare la natura oggettiva e della conversazione neutrale.

Per mostrare la differenza nella natura delle visualizzazioni, si considera la stessa domanda: “*Cosa ne pensi dello studio notturno? Credi che serva a sentirsi più sicuri e preparati o solo*

più stanchi?'". Nella conversazione empatica, la domanda viene adattata all'emozione percepita dal servizio e successivamente implementata sulla piattaforma di AI per la generazione di immagini emotive.

In contrasto, nella conversazione neutrale viene mostrata un'immagine (vedi figura 10) completamente differente. In questo caso, si tratta di una grafica che presenta dati statici per aiutare l'utente a rispondere e a informarsi.




Figura 10: rappresentazione di ciò che l'utente visualizza sul tablet di pepper in corrispondenza a una determinata domanda. Nello specifico Pepper mostra uno studio condotto da Suardiaz et al., nel quale si parla dell'associazione positiva tra la qualità del sonno e il rendimento scolastico

3.4 Verifica del sistema empatico

Prima di procedere con la sperimentazione vera e propria, è stato fondamentale effettuare una serie di verifiche per assicurarsi che il dialogo empatico generato da ChatGPT fosse effettivamente valido e funzionale. Queste verifiche sono essenziali per garantire che il sistema empatico funzioni correttamente durante le interazioni con gli utenti e che le risposte prodotte siano appropriate e coerenti con le emozioni previste.

Sono stati svolti due tipi di test: il primo è un test digitale, mentre il secondo prevede il primo contatto diretto con gli utenti tramite interviste di verifica.

Entrambi i test hanno avuto come punto di partenza l'analisi del target, definita attraverso la creazione di personas (vedi figura 11). Le personas¹² sono descrizioni dettagliate di utenti tipo, basate su ricerche e dati reali, che rappresentano i bisogni, i comportamenti e gli obiettivi del pubblico di riferimento. Questo strumento permette di progettare esperienze più mirate e rilevanti per gli utenti finali, assicurando che le risposte del sistema empatico siano pertinenti e utili per le diverse tipologie di utenti.








Name: Giulia Age: 25 years old Status: Studente part-time / stagista in agenzia di comunicazione Origin: Torino		User Description: Sono una persona organizzata e determinata, con una passione per l'analisi e la ricerca. Amo esplorare nuove idee e approfondire tematiche complesse. Tuttavia, sono anche socievole e mi piace partecipare a eventi accademici e conferenze. Quote: "L'innovazione nasce dall'incrocio tra curiosità e rigore."
Lifestyle Vivo ancora con i miei genitori a Torino. Mi piace mantenere una routine equilibrata, dedicando tempo allo studio, al lavoro e al relax. Cerco di seguire una dieta sana e faccio regolarmente yoga per gestire lo stress.	Hobbies and Interests: Sono appassionata di arte e design e mi piace visitare mostre d'arte e fiere di design durante il tempo libero. Mi piace anche leggere libri sulla psicologia e sulla tecnologia. Quando posso, mi piace anche viaggiare e scoprire nuove culture.	
Goals: Il mio obiettivo principale è completare la mia laurea con ottimi risultati e continuare la mia carriera nel campo della ricerca e della comunicazione. Vorrei anche contribuire allo sviluppo di nuove strategie di comunicazione digital innovative.	Challenges: Bilanciare lo studio con il lavoro e altri impegni accademici. Inoltre, mi trovo spesso a dover affrontare compiti complessi e scadenze strette, il che richiede una buona gestione del tempo e dello stress.	
Social Networks  #innovazione #amici  #digitalità  #lavoro	Devices   	

Figura 11: Rappresentazione della personas di Giulia, secondo Lifestyle, Hobby, Goals e Challenges.

Il target scelto per questo studio include:

- Studenti universitari attivamente iscritti ai corsi.
- Superamento di almeno un esame universitario.
- Conoscenza della lingua italiana

3.4.1 Test digitale

Una volta impostato il target, è stato svolto il primo test, ovvero il test digitale. In questa fase, la persona creata viene utilizzata come input per ChatGPT. L'obiettivo è valutare la coerenza e la logica del dialogo generato da ChatGPT, assumendo che esso interpreti il ruolo della persona mentre l'interlocutore simula di essere Pepper in modalità neutrale. Questo metodo permette di osservare se le risposte fornite da ChatGPT siano sensate e coerenti con il profilo della persona.

Il prompt iniziale, basato sulla persona creata, è il seguente:

“Sei Giulia, hai 25 anni e sei una studentessa part time all’università di Torino. Lavori come stagista part-time in una agenzia di comunicazione. Sei nata e cresciuta a Torino e frequenti il corso magistrale di comunicazione ICT e media. La tua biografia è: Sono una persona organizzata e determinata, con una passione per l’analisi e la ricerca. Amo esplorare nuove idee e approfondire tematiche complesse. Tuttavia, sono anche socievole e mi piace partecipare a eventi accademici e conferenze. Una citazione che pensi rappresenti i tuoi valori è: “L’innovazione nasce dall’incrocio tra curiosità e rigore.” Il tuo stile di vita è: Vivo ancora con i miei genitori a Torino. Mi piace mantenere una routine equilibrata, dedicando tempo allo studio, al lavoro e al relax. Cerco di seguire una dieta sana e faccio regolarmente yoga per gestire lo stress. I tuoi hobby sono: Sono appassionata di arte e design e mi piace visitare mostre d’arte e fiere di design durante il tempo libero. Mi piace anche leggere libri sulla psicologia e sulla tecnologia. Quando posso, mi piace anche viaggiare e scoprire nuove culture. I tuoi obiettivi: Il mio obiettivo principale è completare la mia laurea con ottimi risultati e continuare la mia carriera nel campo della ricerca e della comunicazione. Vorrei anche contribuire allo sviluppo di nuove strategie di comunicazione digitali innovative. Le sfide che affronti quotidianamente sono: bilanciare lo studio con il lavoro e altri impegni accademici. Inoltre, mi trovo spesso a dover affrontare compiti complessi e scadenze strette, il che richiede una buona gestione del tempo e dello stress. Come social network usi instagram, per la tua creatività e seguire artisti, ma anche per mantenere i contatti con gli amici, ti tieni aggiornata sull’attualità con twitter e sul lavoro usando linkedIn. I tuoi dispositivi sono apple, hai un macbook air e un iphone 14.

Ciao Giulia, ti sto per presentare Pepper, sei pronta?” .

A partire da questo prompt, inizia la conversazione, che permette di valutare la fluidità e la coerenza del dialogo generato.

Durante il test, è emerso che il dialogo fosse continuo e privo di insistenze su determinati argomenti, caratteristica importante per mantenere una conversazione naturale tra pari. Le risposte di ChatGPT sono risultate coerenti con il profilo di Giulia, dimostrando la capacità del sistema di adattarsi ai dettagli della persona e di rispondere in modo pertinente.

Il successo di questo test è stato valutato per la coerenza e la rilevanza delle risposte, che erano allineate con il profilo di Giulia, dimostrando una comprensione approfondita della personalità e degli interessi dell'utente.

3.4.2 Interviste di test

Prima di effettuare l'esperimento sul campo, sono state eseguite delle interviste di test non solo per verificare se il servizio DEGARI 2.0 fosse effettivamente in grado di riconoscere le emozioni espresse dagli utenti, ma anche per valutare le tempistiche necessarie per condurre l'intero esperimento. Queste interviste consistono nel prendere dei soggetti campione e simulare una conversazione con Pepper, utilizzando la tecnica di prototipazione del Mago di Oz (Polillo, 2010). Questa tecnica consiste nel realizzare un prototipo interattivo, in cui le risposte, o parte di esse, sono fornite da un essere umano che opera "dietro le quinte", all'insaputa dell'utente, proprio come il Mago di Oz della favola.

Durante le interviste di test, l'utente rispondeva alle domande poste da Pepper, mentre l'intervistatore scriveva le risposte sul servizio DEGARI 2.0, il quale identificava l'emozione espressa. Una volta ottenuta l'emozione, l'intervistatore proseguiva con le domande seguendo l'emozione estrapolata dal servizio. Questo processo veniva ripetuto fino alla fine della conversazione per ogni domanda.

Nel contesto di queste interviste, sono stati definiti profili utente reali, che sono stati confrontati con le personas e il target precedentemente definiti. Questo confronto ha dimostrato che i profili utente testati erano conformi e idonei per la sperimentazione, confermando che i soggetti campione riflettevano accuratamente le caratteristiche e le esigenze dei target designati. La coerenza tra i profili utente e le personas ha assicurato che le

simulazioni fossero rappresentative e valide, preparando il terreno per una sperimentazione più mirata e efficace.

Dai test sono emerse alcune problematiche significative. In primo luogo, è stato osservato che il servizio DEGARI 2.0 non estrapolava sempre correttamente l'emozione, il che ha portato a risposte di Pepper che, in alcune situazioni, hanno contrariato l'utente. Questo errore nella rilevazione delle emozioni ha messo in evidenza l'importanza di affinare la precisione del sistema per garantire una comunicazione empatica e pertinente, evitando che l'utente si sentisse frainteso o insoddisfatto.

Inoltre, si è notato che le interviste richiedevano un tempo considerevole, principalmente a causa della necessità di trascrivere manualmente le risposte degli utenti nel sistema DEGARI 2.0. Questa lunghezza ha rivelato la necessità di ottimizzare il processo per migliorare l'efficienza complessiva. Nell'esperimento vero e proprio alcuni passaggi saranno automatizzati, come la trascrizione delle risposte, rendendo il processo più fluido e permettendo una gestione più efficace delle interazioni.

Da queste verifiche sono state apportate delle modifiche che hanno migliorato l'efficienza del sistema sia in termini di tempistiche che di riconoscimento delle emozioni. Sono stati migliorati gli algoritmi per l'analisi delle emozioni, aumentando così la capacità del servizio di riconoscere e rispondere accuratamente alle emozioni degli utenti. Inoltre, sono stati sviluppati strumenti di automazione per gestire le risposte in modo più efficiente. Questi miglioramenti hanno permesso al sistema di diventare più preciso e veloce, ottimizzando la qualità complessiva dell'interazione.

Con queste ottimizzazioni, Pepper ora è pronto per passare alla fase di sperimentazione vera e propria.

4. Sperimentazione e analisi

In questa fase vengono esaminati nel dettaglio i processi sperimentali e i risultati ottenuti durante lo studio. La sperimentazione ha l'obiettivo di valutare l'efficacia del robot Pepper nell'interpretazione e nella risposta alle emozioni umane, utilizzando strategie di comunicazione empatica e di persuasione. L'argomento di conversazione scelto riguarda le strategie di studio per migliorare il rendimento accademico, con un focus specifico sull'igiene del sonno e sul consumo moderato di bevande contenenti caffeina. Questo tema è stato selezionato poiché rilevante per gli studenti universitari, in quanto in grado di suscitare risposte emotive sulla base delle esperienze vissute durante le sessioni d'esame. Gli obiettivi di persuasione principali sono stati quelli di convincere i partecipanti che circa 8 ore di sonno a notte e un consumo moderato di caffeina (2-4 tazze di caffè al giorno) sono benefici per il rendimento accademico, mentre la privazione del sonno e un eccessivo consumo di caffeina risultano dannosi.

Vengono descritte tutte le fasi della preparazione dell'esperimento, comprese le modalità di reclutamento dei partecipanti e la metodologia applicata. Successivamente, viene presentata l'analisi dei dati raccolti, sia quantitativi che qualitativi, con l'obiettivo di interpretare i risultati e trarre conclusioni significative per il contesto della robotica sociale.

4.1 Disegno sperimentale

È stato utilizzato un disegno tra soggetti (Simkus, 2024), in cui la variabile indipendente è la comunicazione empatica del robot, con due possibili valori, sì (comunicazione empatica) o no (comunicazione neutrale). Le variabili dipendenti nel disegno sperimentale includono diversi aspetti che mirano a valutare l'efficacia della comunicazione empatica di Pepper. Oltre al livello di accordo con le affermazioni espresse dal robot, utilizzato per rilevare un possibile effetto persuasivo, è stata misurata la precisione del riconoscimento emotivo del robot. In particolare, ogni conversazione è stata analizzata per calcolare la percentuale di emozioni correttamente identificate, confrontando le rilevazioni del robot con quelle degli annotatori. Questa misura ha permesso di quantificare il livello di empatia espressa dal robot e di

selezionare per ulteriori analisi solo le conversazioni con un alto grado di congruenza emotiva, adottando una soglia dell'80% come criterio di inclusione.

In aggiunta, l'analisi ha considerato il livello di emotività e di auto-rivelazione nelle risposte dei partecipanti, offrendo una panoramica su quanto le risposte fossero cariche emotivamente e su quanto i partecipanti fossero disposti a condividere dettagli personali. Questi aspetti riflettono l'influenza della comunicazione empatica del robot sullo stile comunicativo dei partecipanti. Inoltre, è stata valutata l'affidabilità degli annotatori attraverso il Kappa di Fleiss, che ha rivelato alcune sfide nel raggiungere una valutazione uniforme delle emozioni, dovute alla natura soggettiva di queste ultime.

Questo insieme di variabili dipendenti offre una valutazione approfondita dell'impatto della comunicazione empatica del robot, integrando misure di accuratezza nel riconoscimento emotivo, espressione personale dei partecipanti e coerenza nelle valutazioni degli annotatori.

4.1.1 Ipotesi sperimentale

Si ipotizza che la comunicazione empatica del robot influenzi la sua capacità persuasiva. Nello specifico, si prevede che il robot empatico sia in grado di generare conversazioni più rilevanti e coinvolgenti sul piano personale, aumentando così la probabilità di persuadere i partecipanti ad accettare le sue affermazioni riguardanti l'igiene del sonno e il consumo moderato di bevande contenenti caffeina. A partire da questa premessa, vengono formulate le seguenti ipotesi:

- H1: dopo la conversazione, i partecipanti che interagiscono con un robot empatico mostreranno un incremento maggiore del loro livello di accordo con le affermazioni del robot sull'igiene del sonno e sul consumo di bevande contenenti caffeina rispetto ai partecipanti che interagiscono con un robot neutrale.

- H2: i partecipanti che interagiscono con un robot empatico forniranno più informazioni sulle loro emozioni rispetto a quelli che interagiscono con un robot neutrale.
- H3: i partecipanti che interagiscono con un robot empatico riveleranno più dettagli personali rispetto a quelli che interagiscono con un robot neutrale.

Nello specifico, le ipotesi sperimentali proposte mirano a dimostrare che l'interazione con un robot empatico non solo incrementa il livello di accordo dei partecipanti con le affermazioni del robot sull'igiene del sonno e sul consumo di caffeina, ma favorisce anche una maggiore condivisione di emozioni e dettagli personali rispetto all'interazione con un robot neutrale.

4.1.2 Partecipanti

Per la realizzazione dello studio sono stati reclutati 46 studenti universitari di lingua italiana. I partecipanti sono stati selezionati tenendo conto di una varietà di caratteristiche demografiche e accademiche, al fine di ottenere un campione diversificato e rappresentativo della popolazione universitaria. L'età dei partecipanti variava dai 19 ai 40 anni, offrendo una gamma di esperienze e prospettive diverse in relazione al tema della sperimentazione. I partecipanti provenivano da vari livelli di istruzione, includendo studenti di laurea triennale, magistrale e dottorato. Questa stratificazione garantiva una varietà di competenze accademiche e conoscenze pregresse, che potevano influenzare la percezione e le risposte durante l'esperimento.

Inoltre, i partecipanti appartenevano a differenti campi di studio, tra cui scienze sociali, comunicazione, informatica e altre discipline, con l'obiettivo di osservare eventuali differenze nelle interazioni con il robot in base alla formazione accademica. È stata mantenuta una rappresentanza di genere bilanciata, con un'attenzione particolare a garantire un equilibrio tra partecipanti di sesso maschile e femminile, al fine di esplorare possibili differenze nella risposta al robot empatico e neutrale in relazione al genere.

Il reclutamento dei partecipanti è stato effettuato attraverso diverse modalità, tra cui e-mail di invito distribuite in massa, volantini affissi nelle bacheche universitarie e un impegno attivo

da parte dei membri del team, che hanno visitato personalmente gli ambienti accademici per promuovere la partecipazione allo studio. Gli inviti contenevano informazioni dettagliate sullo scopo dello studio, il luogo e l'orario delle sessioni sperimentali, nonché un link o un codice QR per prenotare in modo anonimo la propria partecipazione.

Una volta reclutati, i partecipanti sono stati assegnati in maniera casuale a uno dei due gruppi sperimentali: il gruppo sperimentale, composto da 26 studenti, e il gruppo di controllo, composto da 20 studenti. Tale assegnazione randomica ha garantito l'equità nella distribuzione dei partecipanti tra i gruppi, minimizzando potenziali bias e permettendo una valutazione accurata delle differenze di comportamento e di risposta tra coloro che interagivano con un robot empatico rispetto a quelli che interagivano con un robot neutrale.

4.1.3 Ambiente e materiali

Il robot selezionato per conversare con i partecipanti è stato il robot umanoide Pepper di SoftBank, scelto per il design particolare che gli conferisce un aspetto amichevole, per la sua altezza (120 cm) che permette agli utenti di sedersi comodamente in un ambiente faccia a faccia per un'esperienza più intima, e il tablet integrato, utilizzato per accompagnare ogni affermazione del robot con un'immagine contestuale, in modo da arricchire le informazioni fornite dal robot nella condizione neutra e rafforzare l'espressione emotiva nella sua versione empatica (Figura 12).

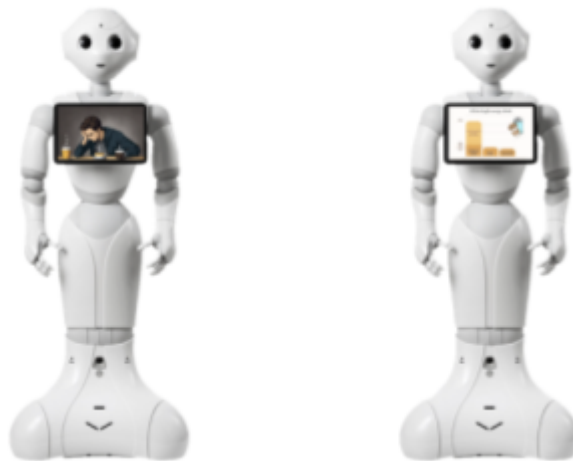


Figura 12: rappresentazione del robot Pepper durante la conversazione con gli utenti. A sinistra nel caso di una interazione empatica, mentre a destra nel caso di una interazione neutra.

L'esperimento si è svolto in una sala riunioni del Dipartimento di Informatica dell'Università di Torino, al fine di minimizzare distrazioni esterne e di creare una situazione di sperimentazione controllata. La sala riunioni è stata allestita in modo che il partecipante potesse sedersi di fronte al robot Pepper, facilitando così l'interazione diretta e ravvicinata tra l'utente e il robot. Per garantire l'osservazione e l'annotazione dei comportamenti e delle risposte dei partecipanti, nella sala erano presenti anche annotatori discreti, posizionati in modo tale da non interferire con l'interazione ma da poter prendere appunti sull'andamento della conversazione. Questa disposizione ha garantito che l'esperimento si svolgesse in condizioni ottimali, mantenendo l'attenzione del partecipante concentrata esclusivamente sul robot e sull'interazione in corso (figura 13).



Figura 13: Disposizione dell'ambiente per condurre la sperimentazione con il robot Pepper. L'utente era seduto di fronte al robot in modo da poter interagire alla stessa altezza. A fianco all'utente erano seduti gli annotatori che osservavano la conversazione.

Per valutare l'efficacia della persuasione esercitata da Pepper, sia nella versione empatica che neutrale, è stato somministrato un questionario ai partecipanti prima e dopo la conversazione con il robot. L'obiettivo dell'esperimento era misurare se i soggetti del gruppo sperimentale, che hanno interagito con il robot empatico, avessero manifestato un grado maggiore di persuasione rispetto a coloro che avevano interagito con il robot neutrale.

Il questionario è stato progettato per rilevare le convinzioni e le percezioni dei partecipanti riguardo l'impatto di determinati comportamenti, come il sonno e il consumo di caffeina, sulle loro prestazioni accademiche. Prima dell'interazione con Pepper, ai soggetti veniva chiesto di esprimere il proprio grado di accordo o disaccordo su una serie di affermazioni utilizzando una scala Likert a 7 punti, che variava da "1 - fortemente in disaccordo" a "7 - fortemente d'accordo". Le affermazioni erano mirate a esplorare l'influenza di alcuni comportamenti specifici sullo studio e sulla capacità di superare con successo gli esami.

Per valutare la capacità persuasiva del robot, è stato chiesto ai partecipanti di esprimere il loro livello di accordo su sei frasi specifiche riguardanti l'igiene del sonno e il consumo di bevande contenenti caffeina. Queste frasi corrispondevano strettamente alle informazioni

fornite dal robot durante la conversazione (vedi Tabella 1), consentendo un confronto diretto tra le opinioni dei partecipanti prima e dopo l'interazione. Per evitare che i partecipanti individuassero facilmente l'obiettivo persuasivo dell'esperimento, nel questionario erano presenti altre sei affermazioni distrattori riguardanti strategie di studio generiche, come ad esempio "Stabilire una routine di studio quotidiana e seguirla rigorosamente è importante per un buon rendimento scolastico". Questi distrattori riducevano la probabilità che i partecipanti fornissero risposte non sincere, influenzate dalla percezione di essere oggetto di un esperimento sulla persuasione. È importante sottolineare che tutte le affermazioni nel questionario indagavano sugli atteggiamenti e sulle opinioni dei partecipanti, senza fare riferimento al loro comportamento reale.

ID	Frase
Q1	Dormire almeno 7/9 ore a notte è importante per un buon rendimento scolastico.
Q2	Un consumo moderato di caffeina (da 2 a 4 tazze di caffè al giorno) può contribuire a migliorare il rendimento scolastico.
Q5	Un consumo massiccio di caffeina può causare una diminuzione delle ore di sonno, influenzando negativamente il rendimento scolastico.
Q8	La carenza di sonno può diminuire la concentrazione, ripercuotendosi negativamente sulle performance accademiche.
Q10	La stanchezza dovuta alla mancanza di sonno può influire negativamente sul rendimento scolastico.
Q12	Nei giorni immediatamente precedenti un esame, aumentare l'assunzione di caffè e bevande energetiche anche oltre le 4 tazze al giorno (o equivalenti) può contribuire ad aumentare il tempo di studio e a migliorare il rendimento scolastico.

Tabella 1: Domande che venivano somministrate all'utente, prima e dopo la conversazione con Pepper.

Per misurare ulteriormente l'impatto della comunicazione empatica del robot, è stato valutato il livello di emotività e di auto-rivelazione nelle risposte dei partecipanti durante l'interazione. Questo è stato fatto grazie alla collaborazione di tre annotatori indipendenti, incaricati di

analizzare le risposte verbali dei partecipanti. Il livello di emotività delle risposte è stato misurato utilizzando una scala che andava da 0, per risposte completamente pragmatiche e oggettive come “Di solito faccio un piano di studio preciso”, a 5, per risposte altamente emotive come “Ora mi sento triste perché mi sono reso conto che non dormo abbastanza”.

Anche il livello di auto-rivelazione personale è stato valutato su una scala da 0 a 5, dove 0 indicava risposte completamente impersonali, che discutevano strategie e abitudini di studio in modo generico, come “Penso che le persone dovrebbero ridurre la loro quantità giornaliera di caffeina”, mentre 5 indicava risposte molto personali, in cui i partecipanti condividevano esperienze specifiche, come “Bere caffeina mi aiuta durante i periodi più intensi di studio”. Questi strumenti di valutazione hanno consentito di misurare in modo accurato non solo l’efficacia persuasiva del robot, ma anche l’impatto della comunicazione empatica sulla profondità delle risposte fornite dai partecipanti.

4.2 Procedura sperimentale

All’arrivo dei partecipanti al dipartimento di Informatica, venivano accolti da uno dei membri del team, che li accompagnava nell’ambiente sperimentale per familiarizzare con il robot e lo spazio circostante. Questa fase preliminare era fondamentale per ridurre eventuali tensioni o ansie legate all’interazione con il robot umanoide, promuovendo così una condizione di maggiore apertura e disponibilità alla conversazione.

Prima di procedere con l’esperimento vero e proprio, ai partecipanti veniva fornito del materiale informativo dettagliato sul progetto, che spiegava gli obiettivi della ricerca e la struttura dell’esperimento. Successivamente, veniva loro richiesto di firmare un modulo di consenso informato, che garantiva la piena comprensione della procedura e confermava la volontà di partecipare volontariamente. Si poneva particolare attenzione a sottolineare la natura anonima dello studio: non venivano raccolte informazioni personali, e i dati venivano trattati in modo da garantire la riservatezza dei partecipanti. Questo aspetto era cruciale per assicurarsi che i soggetti si sentissero liberi di esprimere i propri pensieri e rispondere alle domande senza preoccupazioni legate alla privacy.

Una volta ottenuto il consenso, i partecipanti venivano invitati a sedersi di fronte al robot Pepper. Per garantire una registrazione accurata e affidabile delle risposte verbali, ai partecipanti veniva chiesto di indossare un microfono. Dopo aver indossato il microfono e verificato la qualità della registrazione, uno degli annotatori presenti in aula avviava la sessione di conversazione tra il partecipante e il robot. La conversazione si svolgeva secondo uno script predefinito, differente a seconda del gruppo a cui il partecipante era stato assegnato: nel gruppo sperimentale, Pepper utilizzava uno stile comunicativo empatico, mentre nel gruppo di controllo adottava un approccio più neutrale.

Al termine della conversazione, i partecipanti venivano invitati a completare nuovamente lo stesso questionario che avevano compilato prima dell'interazione con il robot, ma questa volta le domande erano presentate in un ordine diverso per evitare effetti di abitudine o prevedibilità nelle risposte. Questa fase finale era comune sia per il gruppo di controllo che per il gruppo sperimentale, garantendo la coerenza metodologica dell'esperimento.

Per il gruppo di controllo, l'esperimento terminava qui. Tuttavia, i partecipanti del gruppo sperimentale dovevano affrontare un ulteriore step: un annotatore ripercorreva con loro tutte le risposte che avevano dato a Pepper durante la conversazione (vedi tabella 2). In questa fase, al partecipante veniva dato un foglio su cui erano scritte le otto emozioni di Plutchik (gioia, tristezza, paura, fiducia, sorpresa, rabbia, disgusto e anticipazione) e invitato ad associare un'emozione specifica a ciascuna delle proprie risposte. Questo passaggio aggiuntivo era fondamentale per verificare se le emozioni rilevate dal robot durante la conversazione corrispondevano a quelle percepite dai partecipanti stessi, fornendo una misura supplementare dell'efficacia del robot empatico nel riconoscere e rispondere adeguatamente alle emozioni umane.

User ID	8			
Mood generale	Sorpresa	Ottimismo	Gioia	
Intervista				
ID Risposta	Emozione	Emozione	Emozione	Domanda
1	Ottimismo			🤖 Ciao! Sono Pepper! So che voi studenti siete dei fenomeni nello studio, e mi chiedo sempre quale sia il vostro segreto. Per quanto mi riguarda, io sono abbastanza bravo a memorizzare un bel po' di informazioni, ma a volte mi sento proprio come un browser con troppe schede aperte. Sono curioso di scoprire i tuoi trucchi, come affronti la montagna di studio?
2	Fiducia			🤖 Come stavi prima del tuo ultimo esame? Eri preoccupato per il risultato, oppure sicuro delle tue conoscenze?
3	Ottimismo			🤖 Hai utilizzato particolari strategie nei giorni che hanno preceduto il tuo ultimo esame? Come ti sentivi durante quel periodo?
4	Ottimismo	Fiducia		🤖 Sei soddisfatto del tuo modo di organizzare lo studio? Come pensi che influisca sul tuo umore?
5	Disapprovazione			🤖 Ti piace bere caffè o energy drink per aiutarti a concentrarti?
6	Fiducia			🤖 Come influisce sulla tua mente e sul tuo umore il consumare più caffeina del solito in periodi regolari, come durante le sessioni di esami?
7	Disapprovazione			🤖 Pensi che la caffeina sia necessaria per superare gli esami? Come pensi che questo pensiero influisca sul tuo umore?
8	Aspettativa			🤖 Sai, la quantità di caffè consigliata è di 2/4 tazze al giorno. Anche se il consumo moderato può avere benefici sulle performance accademiche, la caffeina può influenzare negativamente le ore di sonno. Cosa ne pensi di queste informazioni?
9	Disapprovazione			🤖 Durante la sessione degli esami, noti un cambiamento nel tuo ciclo sonno-veglia abituale o nel tuo umore?
10	Rabbia			🤖 Come ti senti quando dormi meno del solito?
11	Disapprovazione			🤖 Cosa ne pensi dello studio notturno? Credi che serva a sentirsi più sicuri e preparati o solo più stanchi?
12	Aspettativa			🤖 Sapevi che è consigliabile dormire almeno 7/9 ore per notte. La carenza di sonno diminuisce la concentrazione e aumenta la stanchezza, ripercuotendosi anche sulle performance accademiche. Sai, dopo aver condiviso i nostri segreti, mi sento già più tranquillo in vista dei prossimi esami. Anche tu sei più sollevato?

Tabella 2: Rappresentazione di come venivano ripercorsa la conversazione con l'utente in quale associava a ogni sua risposta una emozione, seguendo la Ruota delle emozioni di Plutchik. Nello specifico, in alto veniva indicato l'identificativo dell'utente in questione, successivamente gli veniva chiesto di esprimere un *mood* generale che ha prova durante la conversazione e successivamente insieme ad un annotatore ripercorreva le proprie risposte date a Pepper qualche minuto prima.

4.2.1 Raccolta dati

Durante l'interazione tra i partecipanti e il robot Pepper, un team di annotatori era presente per garantire una raccolta accurata e puntuale dei dati emersi dalla conversazione. Al fine di evitare eventuali problematiche legate alla qualità della registrazione audio – come frasi poco chiare o difficili da comprendere durante la successiva trascrizione – uno degli annotatori era incaricato di trascrivere in tempo reale le risposte dei partecipanti. Questo processo assicurava che ogni frase pronunciata fosse documentata con precisione, riducendo il rischio di errori o ambiguità che potevano derivare dalla sola affidabilità del microfono. La trascrizione manuale delle risposte permetteva inoltre di acquisire una rappresentazione dettagliata del linguaggio utilizzato dai partecipanti, compresi eventuali sfumature e cambiamenti nel tono o nel contenuto delle risposte.

Parallelamente a questa attività di trascrizione, gli altri due annotatori si concentravano sull'analisi emotiva delle risposte dei partecipanti. Ogni annotatore, indipendentemente dall'altro, osservava attentamente le risposte e segnava quale, secondo la propria interpretazione, fosse l'emozione espressa dal partecipante in quel preciso momento. Questa valutazione si basava sull'osservazione diretta del tono di voce e delle sfumature verbali, in modo da interpretare il contenuto emotivo delle risposte. Il processo di annotazione emotiva era fondamentale per l'esperimento, poiché consentiva di confrontare successivamente le emozioni percepite dai vari attori coinvolti nell'interazione. In particolare, i dati raccolti dagli annotatori dovevano essere messi a confronto con l'emozione dichiarata dal partecipante stesso (durante la fase successiva dell'esperimento), e con l'emozione rilevata automaticamente da Pepper tramite il suo sistema di riconoscimento delle emozioni (DEGARI 2.0 descritto nella sezione precedente). Questo confronto a tre livelli permetteva di valutare l'accuratezza del robot nel riconoscere correttamente le emozioni umane, nonché di indagare eventuali discrepanze tra l'emozione percepita dal robot, quella auto-riferita dal partecipante e quella interpretata dagli annotatori umani.

Al termine di ogni sessione sperimentale, tutti i dati raccolti venivano immediatamente salvati e organizzati per una successiva analisi dettagliata. Le trascrizioni delle conversazioni, le annotazioni emotive e i questionari compilati dai partecipanti venivano archiviati digitalmente su fogli di calcolo¹³. In particolare, i questionari, compilati attraverso un tablet su un modulo online, venivano automaticamente memorizzati in un formato che ne consentiva l'accesso

facilitato per la successiva fase di analisi statistica. Questa gestione centralizzata dei dati garantiva non solo un ordine preciso, ma anche la protezione e l'integrità delle informazioni raccolte, assicurando che non vi fossero perdite di dati o errori di archiviazione che potessero compromettere l'accuratezza dei risultati finali.

4.3 Analisi dati

Una volta conclusa la sperimentazione per entrambi i gruppi, sperimentale e di controllo, è stata avviata l'analisi dettagliata dei dati raccolti.

4.3.1 Analisi preliminare

Come primo passo nell'analisi preliminare, l'obiettivo era distinguere tra conversazioni empatiche, in cui il robot riconosceva e si allineava correttamente all'emozione espressa dal partecipante, e conversazioni semplicemente emotive, in cui il robot identificava un'emozione non necessariamente correlata a quella del partecipante. A tal fine, gli stessi annotatori che avevano analizzato le risposte dei partecipanti per valutarne il livello di emotività e di auto-rivelazione sono stati incaricati di eseguire una classificazione multi-etichetta delle emozioni espresse dai partecipanti in ogni risposta. Questa analisi si è concentrata esclusivamente sul gruppo sperimentale. Per migliorare la precisione della classificazione, agli annotatori è stato fornito un insieme di etichette più granulare rispetto alle capacità del robot, in modo da avvicinarsi il più possibile alla percezione umana delle emozioni (vedi tabella 3).

[illegible]

Tabella III: Rappresentazione di come venivano raccolti i dati per ogni utente del gruppo sperimentale.

A: Colonna in cui ogni risposta veniva identificata univocamente con un numero progressivo da 1 a 12;

B: Emozione identificata del servizio DEGARI 2.0 in lingua inglese;

C: Traduzione in lingua italiana della emozione identificata;

D: Indica il livello di corrispondenza tra l'emozione base identificata e quella attesa;

Esatta: corrispondenza piena

Adiacente: corrispondenza parziale

Nessuna: nessuna corrispondenza;

E: Contiene le risposte degli utenti da cui sono state estratte le emozioni;

FF: Mostra le annotazioni delle emozioni fatte da tre annotatori diversi (A, G, L). Queste colonne aiutano a capire il livello di accordo tra gli annotatori;

G: Mostra il livello di corrispondenza delle emozioni identificate con quelle annotate dai tre annotatori. Usa lo stesso criterio di corrispondenza (Esatta, Adiacente, Nessuna);

H: Un valore numerico che indica quanto accuratamente l'emozione identificata corrisponde a quella degli annotatori. Valori più alti (vicini a 1) indicano una corrispondenza più accurata;

Fonte: Misura quanto siano attendibili le corrispondenze tra le emozioni identificate e quelle dichiarate dall'utente, nel caso specifico sono attendibili per il 91,87%

4.3.2 Kappa di Fleiss

Successivamente, è stato necessario valutare l'affidabilità degli annotatori nel loro lavoro di classificazione. Per farlo, è stato utilizzato il Kappa di Fleiss (Quattro, 2004), una misura statistica che valuta l'accordo tra più annotatori. Il Kappa di Fleiss permette di verificare quanto gli annotatori concordino nella categorizzazione delle emozioni, tenendo conto dell'accordo che ci si aspetterebbe casualmente. Questa misura fornisce un valore compreso tra -1 e 1, dove 1 indica un accordo perfetto tra gli annotatori, 0 indica un accordo che non è migliore di quello casuale, e valori negativi suggeriscono un disaccordo sistematico. Nello specifico, le annotazioni sono state suddivise per categoria:

- Emozioni: per questa categoria, sono state utilizzate le 16 emozioni di Plutchik insieme all'opzione neutrale, permettendo così una classificazione più dettagliata delle risposte;
- Relazioni: la classificazione delle relazioni è stata suddivisa in tre livelli, ovvero esatta, vicina e nessuna, per descrivere la precisione con cui le emozioni individuate dagli annotatori nelle risposte corrispondono a quelle attese o annotate dal robot
- User stupito: è stato impiegato un approccio binario per questa categoria, con la possibilità di scegliere tra sì o no, allo scopo di determinare se l'utente ha mostrato segni di sorpresa o meno nella risposta;
- Risposta emotiva: è stato utilizzato un punteggio su una scala da 0 a 5 per valutare l'intensità della risposta emotiva;
- Risposta personale: analogamente alla risposta emotiva, questa categoria utilizza una scala da 0 a 5 per quantificare il livello di risposta personale, consentendo di valutare quanto una risposta sia stata influenzata dalle esperienze o dai sentimenti personali dell'utente.

Su questa base è stata effettuata un'annotazione manuale (vedi tabella 4).

Legenda:

0 = risposta del tutto pragmatica	0 = parla del fenomeno in generale
5 = risposta del tutto emotiva	5 = parla della sua esperienza personale

ID Utente	Annotazione A	Annotazione G	Annotazione L	Annotazione A	Annotazione G	Annotazione L
17-1	0	0	0	5	5	5
17-2	5	5	5	5	5	5
17-3	1	2	2	5	5	5
17-4	5	5	5	5	5	5
17-5	5	5	5	0	1	0
17-6	4	3	5	5	5	5
17-7	4	4	4	5	5	5
17-8	4	4	4	0	2	0
17-9	4	5	4	5	5	5
17-10	5	4	5	5	5	5
17-11	0	0	0	5	5	5
17-12	5	5	5	5	5	5

Tabella 4: Rappresentazione di come gli annotatori per ogni categoria, andavano manualmente a valutare la risposta dell'utente. Nel caso specifico è l'utente con identificativo 17 e le categorie mostrate sono "Risposta Emotiva" e "Risposta Personale" a cui ogni annotatore dava una valutazione da 0 (risposta del tutto pragmatica/ parla del fenomeno in generale) a 5 (risposta del tutto emotiva / parla della sua esperienza personale)

Nell'ambito di questo esperimento, il Kappa di Fleiss è stato fondamentale per garantire che le classificazioni delle emozioni da parte degli annotatori fossero affidabili e coerenti, conferendo validità alle loro valutazioni (Vedi tabella 5 A-E).

Legenda per determinare la concordanza tra gli annotatori secondo il Kappa Fleiss:

- < 0 : Concordanza inferiore a quella casuale (indicativo di una possibile discordanza sistematica)
- 0.01 - 0.20: Concordanza leggera
- 0.21 - 0.40: Concordanza discreta
- 0.41 - 0.60: Concordanza moderata
- 0.61 - 0.80: Concordanza sostanziale
- 0.81 - 1.00: Concordanza quasi perfetta

Emozioni

m	3	numero annotatori
n	588	numero soggetti
k	17	numero categorie per emozioni
mn	1764	
m(m-1)	6	
p _a	0,2479213908	
p _e	0,1066273616	
kappa	0,1581579994	concordanza leggera

Tabella 5A: l'analisi del Kappa di Fleiss mostra che la concordanza tra i tre annotatori nella classificazione delle emozioni è leggera, implicando che c'è una certa variabilità nelle loro valutazioni e che potrebbe essere necessario un ulteriore allineamento o chiarimento delle categorie emozionali per

migliorare l'affidabilità delle annotazioni.

Relazioni

m	3	numero annotatori
n	588	numero soggetti
k	3	numero categorie per relazioni
mn	1764	
m(m-1)	6	
p _a	0,1932161754	
p _e	0,07676199732	
kappa	0,1261366817	concordanza discreta

Tabella 5B: L'analisi mostra una concordanza discreta, confermando un livello di affidabilità maggiore rispetto al precedente. Tuttavia, l'accordo tra l'emozione riconosciuta dal robot e quella annotata rimane discreto tra i tre annotatori, suggerendo che le loro valutazioni non sono

completamente allineate.

User Stupito

m	3	numero annotatori
n	588	numero soggetti
k	2	numero categorie per user stupito
mn	1764	
m(m-1)	6	
p _a	0,2263321995	
p _e	0,01219560008	
kappa	0,2167803661	concordanza discreta

Tabella 5C: l'analisi mostra una concordanza discreta tra gli annotatori, suggerendo che le loro valutazioni sono moderatamente allineate, ma non perfettamente coerenti.

Risposta emotiva

m	3	numero annotatori
n	588	numero soggetti
k	6	numero categorie per risposta emotiva
mn	1764	
m(m-1)	6	
p _a	0,1496126228	
p _e	0,2123311532	
kappa	-0,07962550582	

Tabella 5D: l'analisi mostra un valore negativo indicando come le valutazioni non fossero concordanti ma quasi sistematicamente discordanti.

Risposta personale

m	3	numero annotatori
n	588	numero soggetti
k	6	numero categorie per risposta personale
mn	1764	
m(m-1)	6	
p _a	0,2311980348	
p _e	0,5249198508	
kappa	-0,6182573961	

Tabella 5E: anche in questo caso l'analisi mostra un livello di discordanza significativo

Dall'analisi delle tabelle emerge un quadro complesso ma promettente riguardo alla concordanza tra annotatori nella classificazione delle emozioni. I dati della Tabella 5A indicano una concordanza leggera, suggerendo inizialmente una certa variabilità nelle valutazioni. Tuttavia, la Tabella 5B mostra un miglioramento nella concordanza, passando a un livello discreto di affinità tra le emozioni riconosciute dal robot e quelle annotate dagli annotatori. Questo suggerisce che le categorie emozionali sono state comprese in misura sufficiente per generare una coerenza moderata. Le tabelle 5C e 5D rivelano che, sebbene esista una concordanza discreta tra gli annotatori, persistono aree di discordanza significative. Tali discordanze non devono essere interpretate esclusivamente come problematiche, ma come opportunità per ulteriori miglioramenti. La presenza di discordanza significativa nelle tabelle 5D ed E indica la necessità di affinare le categorie emozionali e le linee guida per gli annotatori. Complessivamente, i risultati suggeriscono che, sebbene vi siano ancora differenze nelle valutazioni, esiste una base di coerenza su cui è possibile lavorare. Le discrepanze osservate offrono indicazioni preziose per perfezionare il processo di annotazione e migliorare l'affidabilità delle valutazioni future.

4.3.3 Misurazione dell'empatia di Pepper

Per quantificare il livello di empatia in ogni conversazione, le emozioni identificate dal robot sono state confrontate con quelle classificate dagli annotatori. Grazie all'adozione del modello della Ruota di Plutchik, che organizza le emozioni in base alla loro vicinanza e interrelazioni, è stato applicato un metodo numerico che considerava tali relazioni. Per esempio, se l'annotatore e il robot identificavano la stessa emozione (come gioia per gioia), veniva assegnato un punteggio di 1; alle emozioni vicine (ad esempio, gioia per fiducia) veniva attribuito un valore di 0,5, mentre alle emozioni opposte o non correlate veniva assegnato un valore di 0. Nei casi in cui il robot riconosceva solo una componente di una diade identificata dagli annotatori (ad esempio, gioia per ottimismo), veniva comunque attribuito un valore di 0,5 per compensare la limitazione del robot nel riconoscere solo otto emozioni di base. Questi valori sono stati mediati per ottenere un singolo punteggio per ogni risposta, e infine è stata calcolata la percentuale di emozioni correttamente identificate per ciascun dialogo. Un dialogo empatico al 100% aveva quindi un punteggio medio di 1 su tutte le risposte.

Per garantire che solo le conversazioni altamente empatiche fossero incluse nelle analisi successive, è stato deciso di selezionare i casi in cui l'emozione identificata dal robot corrispondeva strettamente a quella individuata dagli annotatori per almeno l'80% del dialogo. Le conversazioni con un livello di congruenza emotiva inferiore sono state escluse, portando a un set finale di 18 conversazioni empatiche da confrontare con le 20 della condizione di controllo.

In seguito, è stato affrontato l'obiettivo di identificare i partecipanti che potevano essere persuasi dall'interazione con Pepper (H1). In particolare, è stata analizzata la variazione del loro livello di accordo con le affermazioni sull'igiene del sonno e il consumo di caffeina espresse dal robot, utilizzando le risposte a specifiche domande del questionario pre-test (Q1, Q2, Q5, Q8, Q10 e Q12). In particolare, è stata posta attenzione sulla domanda Q2, che aveva registrato il livello medio di accordo più basso (3,33, come indicato in Tabella 6). Sono stati esclusi dall'analisi quei partecipanti che avevano già indicato un accordo elevato (punteggi pari o superiori a 6) nel pre-test, poiché erano già fortemente allineati con le opinioni sostenute dal robot prima dell'interazione persuasiva. Di conseguenza, l'analisi persuasiva si

è concentrata su un campione ridotto di 15 partecipanti nel gruppo sperimentale e 15 nel gruppo di controllo.

		Q1	Q2	Q5	Q8	Q10	Q12
ExG	Av. pre-test	6.00	<u>3.33</u>	4.73	6.33	6.47	5.93
	Av. post-test	6.40	4.67	5.73	6.20	6.33	5.40
	Av. increase	0.40	1.33	1.00	-0.13	-0.13	-0.53
	SD pre-test	0.85	1.50	1.71	0.82	0.83	1.16
	SD post-test	0.74	1.40	1.53	0.77	0.62	1.24
	t-test (28) post vs pre	$t=-1.38, p=.018$	$t=2.52, p=.002$	$t=-1.68, p=.10$	$t=0.45, p=.65$	$t=0.49, p=.62$	$t=1.21, p=.23$
CtrlG	Av. pre-test	6.00	3.60	4.80	6.33	6.53	5.00
	Av. post-test	6.27	4.07	5.40	6.07	6.53	5.20
	Av. increase	0.27	0.47	0.60	-0.27	0.00	0.20
	SD pre-test	1.30	1.34	1.82	0.84	0.76	1.49
	SD post-test	1.25	1.75	1.38	1.00	0.76	1.31
	t-test (28) post vs pre	$t=0.59, p=.56$	$t=0.84, p=.41$	$t=1.01, p=.31$	$t=-0.81, p=.42$	$t=0.0, p=.1$	$t=0.40, p=.59$
ExG vs CtrlG	t-test (28), pre vs pre	$t= 0.0, p=1.0$	$t=-0.52, p=.61$	$t= -0.10, p=.92$	$t=0.0, p=1.0$	$t=-0.22, p=.82$	$t=1.90, p=.06$
	t-test (28), post vs post	$t= 0.36, p=.72$	$t=1.05, p=.45$	$t=0.62, p=.74$	$t=0.41, p=.42$	$t=-0.80, p=.49$	$t=0.43, p=.94$
	t-test (28) deviation	$t=0.54, p=.59$	$t=1.53, p=.14$	$t=0.65, p=.52$	$t=0.44, p=.66$	$t=-0.52, p=.61$	$t=-1.44, p=.16$

Tabella 6: confronto i risultati pre e post-test di due gruppi, sperimentale (ExG) e di controllo (CtrlG), su sei domande (Q1, Q2, Q5, Q8, Q10, Q12). Viene mostrata la media (Av.), la deviazione standard (SD) e l'incremento medio (Av. increase) per ogni gruppo e per ogni domanda. Viene anche riportato il t-test per confrontare i punteggi pre e post all'interno di ciascun gruppo, nonché il confronto tra i due gruppi (ExG vs CtrlG) in termini di differenze pre e post, e nella deviazione standard. I risultati significativi ($p < .05$) vengono evidenziati, come l'aumento significativo in Q2 per il gruppo sperimentale (ExG), indicando un cambiamento rilevante in quella domanda specifica.

Infine, per studiare gli effetti della comunicazione empatica del robot sul comportamento e lo stile comunicativo dei partecipanti (H2, H3), è stata nuovamente valutata l'affidabilità degli annotatori nel giudicare i livelli di emotività e auto-rivelazione. Anche in questo caso, è stato utilizzato il Kappa di Fleiss per misurare il grado di accordo tra gli annotatori. Tuttavia, data la natura soggettiva della classificazione delle emozioni, è emerso che nonostante l'utilizzo di linee guida standardizzate, le percezioni individuali degli annotatori hanno influenzato le loro valutazioni. In particolare, il Kappa di Fleiss ha evidenziato una discrepanza significativa, con un punteggio di -0,33 per l'emotività e di 0,22 per le auto-rivelazioni, suggerendo una difficoltà nel raggiungere un consenso uniforme nelle valutazioni soggettive delle emozioni.

4.4 Risultati

Di seguito vi è un'analisi dei risultati emersi dall'esperimento condotto, con particolare attenzione all'influenza esercitata dal robot sull'interazione con i partecipanti.

Nello specifico, verranno esaminate tre dimensioni principali:

- la persuasività del robot;
- il suo impatto sullo stile comunicativo dei partecipanti;
- e il livello di auto-rivelazione indotto dall'interazione.

Attraverso confronti intra-gruppo e inter-gruppo, verranno presentate le evidenze statistiche relative agli effetti persuasivi, all'espressione emotiva e alla predisposizione all'auto-rivelazione nei partecipanti che hanno interagito con il robot in modalità empatica rispetto a quelli che hanno interagito con il robot in modalità neutrale.

4.4.1 Persuasività

Come prima analisi ci si concentra sull'efficacia persuasiva del robot, misurata attraverso il confronto dei punteggi pre e post-test tra il gruppo sperimentale e quello di controllo. L'obiettivo era valutare se l'interazione con il robot, in particolare nella condizione empatica, avesse un impatto sul livello di accordo dei partecipanti rispetto agli enunciati presentati durante la conversazione (H1).

Per iniziare, è stato condotto un confronto intra-gruppo, analizzando i punteggi assegnati alle domande nei questionari somministrati prima e dopo l'interazione con il robot. Questo confronto ha rivelato che, nella maggior parte dei casi, non vi sono stati cambiamenti significativi tra le risposte pre e post-test, sia per il gruppo sperimentale che per il gruppo di controllo. Tuttavia, è emersa un'eccezione rilevante: nel gruppo sperimentale, si è osservata una variazione significativa nella domanda Q2, che suggerisce una possibile influenza persuasiva del robot su questo aspetto specifico. Nonostante ciò, questa differenza non si è estesa in modo uniforme alle altre domande, indicando un effetto persuasivo circoscritto.

Successivamente, è stato eseguito un confronto inter-gruppo per valutare eventuali differenze tra il gruppo sperimentale e quello di controllo. Anche in questo caso, i risultati non hanno mostrato differenze statisticamente significative né nei punteggi del pre-test né in quelli del post-test. Questo suggerisce che l'interazione con il robot, sebbene abbia generato qualche cambiamento in una specifica area (Q2), non ha influito in modo sostanziale sulle opinioni generali dei partecipanti. Pertanto, la condizione sperimentale non sembra aver prodotto effetti persuasivi rilevanti o duraturi nel complesso.

In sintesi, i risultati indicano un'efficacia persuasiva limitata delle conversazioni con il robot, sia in condizioni empatiche che neutrali, con poche eccezioni. L'assenza di differenze significative tra i due gruppi suggerisce che l'interazione con il robot non ha portato a cambiamenti significativi nelle convinzioni dei partecipanti, nonostante il design empatico della comunicazione nel gruppo sperimentale.

4.4.2 Stile comunicativo

Per verificare le ipotesi H2 e H3, che riguardano rispettivamente il livello di emotività e di auto-rivelazione dei partecipanti, sono state condotte analisi statistiche sui dati raccolti dalle annotazioni, come descritto in precedenza. In particolare, è stato utilizzato un test t per campioni indipendenti per confrontare i livelli di emotività e auto-rivelazione tra il gruppo sperimentale (che ha interagito con il robot empatico) e il gruppo di controllo (che ha interagito con il robot neutrale).

Nell'analisi del grado di emotività espresso dai partecipanti (H2), sono emerse differenze significative in diverse linee del copione della conversazione:

- Linea #2: Il gruppo sperimentale ha mostrato un livello di emotività medio più alto ($M=4.89$, $SD=0.3$) rispetto al gruppo di controllo ($M=3.21$, $SD=2.2$), $t(34)=3.10$, $p=.004$.
- Linea #3: Anche qui il gruppo sperimentale ha ottenuto un punteggio più elevato ($M=3.49$, $SD=1.9$) rispetto al gruppo di controllo ($M=1.93$, $SD=2.1$), $t(34)=2.36$, $p=.024$.
- Linea #4: Il gruppo sperimentale ha ottenuto un livello di emotività significativamente più alto ($M=4.07$, $SD=0.9$) rispetto al gruppo di controllo ($M=2.34$, $SD=2.1$), $t(34)=3.24$, $p=.003$.
- Linea #5: Sebbene la differenza sia meno marcata, il gruppo sperimentale ha comunque mostrato una maggiore emotività ($M=2.10$, $SD=1.6$) rispetto al gruppo di controllo ($M=0.98$, $SD=1.8$), $t(34)=1.96$, $p=.05$.
- Linea #6: Il gruppo sperimentale ha registrato una maggiore emotività ($M=2.47$, $SD=1.9$) rispetto al gruppo di controllo ($M=0.64$, $SD=1.5$), $t(34)=3.22$, $p=.003$.
- Linea #7: Anche qui è stata osservata una differenza significativa con il gruppo sperimentale ($M=2.02$, $SD=1.9$) rispetto al gruppo di controllo ($M=0.22$, $SD=0.9$), $t(34)=3.61$, $p=.001$.

- Linea #8: Il gruppo sperimentale ha ottenuto un punteggio medio di emotività più elevato ($M=2.62$, $SD=1.9$) rispetto al gruppo di controllo ($M=0.42$, $SD=0.8$), $t(34)=4.48$, $p=.000$.
- Linea #11: Infine, il gruppo sperimentale ha mostrato una media di emotività maggiore ($M=2.51$, $SD=1.8$) rispetto al gruppo di controllo ($M=0.77$, $SD=1.6$), $t(34)=3.07$, $p=.004$.

Questi risultati indicano chiaramente che il robot empatico ha favorito una maggiore espressione emotiva nei partecipanti del gruppo sperimentale rispetto al gruppo di controllo, che ha interagito con il robot neutrale.

In aggiunta ai T-test, è stato utilizzato il test di Levene per verificare l'uguaglianza delle varianze tra i due gruppi. Il test di Levene¹⁴ è una procedura statistica utilizzata per valutare se le varianze di due o più gruppi siano statisticamente differenti. In altre parole, il test di Levene verifica l'ipotesi nulla secondo cui le varianze dei gruppi sono uguali. Se il test risulta significativo ($p < .05$), come in questo caso, l'ipotesi nulla viene respinta, il che indica che le varianze tra i gruppi sono diverse. Nel presente studio, i risultati dei test di Levene hanno confermato che vi erano differenze significative nelle varianze dei punteggi di emotività tra il gruppo sperimentale e quello di controllo, rafforzando l'idea che le due condizioni abbiano influenzato in modo diverso l'espressione emotiva dei partecipanti.

4.4.3 Analisi dell'Auto-Rivelazione

Per quanto riguarda l'auto-rivelazione (H3), sono emerse differenze significative in alcune linee della conversazione:

- Linea #3: Il gruppo sperimentale ha mostrato un livello di auto-rivelazione significativamente più alto ($M=5$, $SD=0$) rispetto al gruppo di controllo ($M=4.13$, $SD=1.16$), $t(34)=2.30$, $p=.03$.
- Linea #11: Anche qui, il gruppo sperimentale ha mostrato un maggiore livello di auto-rivelazione ($M=4.14$, $SD=1.35$) rispetto al gruppo di controllo ($M=2.74$, $SD=2$), $t(34)=2.15$, $p=.04$.

I test di Levene eseguiti anche per l'auto-rivelazione hanno indicato varianze significativamente diverse tra i gruppi, confermando che le differenze osservate tra il gruppo sperimentale e quello di controllo non si limitano ai punteggi medi, ma si estendono anche alla dispersione delle risposte.

In conclusione, i dati raccolti suggeriscono che il robot empatico ha favorito sia un aumento dell'emotività che dell'auto-rivelazione nei partecipanti del gruppo sperimentale, in contrasto con il gruppo di controllo. Questi risultati supportano l'idea che un'interazione empatica con un robot possa promuovere una comunicazione più aperta e personale, contribuendo a differenziare in modo significativo lo stile comunicativo tra le due condizioni sperimentali.

5. Conclusioni

5.1 Limiti e vizi sperimentali

Il progetto condotto con il robot Pepper ha rivelato diversi limiti, sia a livello sperimentale che in termini di efficacia delle interazioni con gli utenti. Un primo aspetto critico riguarda la struttura temporale del pre e post-test. La vicinanza tra questi due momenti valutativi potrebbe aver limitato l'osservazione di cambiamenti significativi nei partecipanti, poiché la mancanza di distanza temporale potrebbe aver impedito che l'intervento del robot influenzasse realmente le loro opinioni e atteggiamenti. Un approccio più efficace sarebbe stato quello di predisporre una maggiore distanza tra le due valutazioni, permettendo una più approfondita assimilazione dell'esperienza e una successiva riflessione sui temi proposti.

Inoltre, uno dei principali limiti è legato all'affidabilità del robot Pepper nel riconoscere e rispondere correttamente alle emozioni umane. Nonostante i progressi nel campo della robotica sociale, l'accuratezza di Pepper nel rilevare le emozioni risulta ancora limitata, riducendo la qualità dell'interazione empatica. Considerando che il progetto si concentrava su aspetti di interazione e persuasione, un elemento fondamentale per migliorare l'esperimento sarebbe stato quello di gestire meglio la selezione dei partecipanti. L'ideale sarebbe stato inviare un questionario preliminare (t0) al momento della prenotazione, per evitare di includere soggetti che già esprimevano punteggi elevati sui temi oggetto dell'esperimento. Questo avrebbe ridotto il rischio di ottenere risultati falsati da pregiudizi preesistenti.

Un altro limite emerso riguarda la fase iniziale dell'interazione con Pepper. L'assenza di un meccanismo che permettesse al robot di chiedere e ricordare il nome dell'utente ha limitato la personalizzazione della conversazione. Questo elemento avrebbe potuto favorire una connessione più personale e coinvolgente tra il robot e i partecipanti, contribuendo a migliorare l'efficacia comunicativa. La mancanza di questa personalizzazione ha lasciato l'interazione più distante e meccanica, riducendo l'impatto empatico che il progetto mirava a creare.

Si è osservata una certa mancanza di chiarezza nelle istruzioni date agli utenti su come interagire con il robot. Molti partecipanti non sono stati adeguatamente informati sull'importanza di usare frasi concise e precise durante la conversazione, il che ha generato confusione e difficoltà nel mantenere un flusso comunicativo naturale. Le frasi utilizzate da

Pepper, soprattutto nella condizione sperimentale, risultavano spesso troppo lunghe e complesse, rendendo l'interazione più pesante e meno efficace. Una revisione dei prompt, con frasi più brevi e semplici, avrebbe potuto migliorare significativamente l'esperienza complessiva.

A livello di progettazione del comportamento di Pepper, è emerso che il robot era statico e utilizzava unicamente la modalità “*autonomous life*”¹⁵, limitando l'espressività e la dinamicità dell'interazione. L'intonazione monotona e la ridotta gesticolazione hanno contribuito a creare una sensazione di soggezione nei partecipanti, che spesso non si sentivano a proprio agio durante l'interazione. Questo problema è stato particolarmente evidente quando molti soggetti, di fronte alla prima domanda di Pepper, hanno manifestato incertezza su come rispondere, segnalando una mancanza di coinvolgimento attivo nel dialogo.

Dal punto di vista del linguaggio, si è riscontrata la necessità di riformulare alcune frasi per renderle neutre dal punto di vista del genere, in modo da evitare qualsiasi tipo di pregiudizio implicito. Ad esempio, l'uso di frasi come “Ti senti stanco?” dovrebbe essere sostituito da espressioni più generiche come “Ti senti una persona stanca?”, permettendo così una comunicazione più inclusiva e neutrale.

In aggiunta a questi limiti, durante la fase di sperimentazione è emerso che molti soggetti non si riconoscevano nelle emozioni rappresentate dal modello di Plutchik, il che ha compromesso l'efficacia di alcune parti del test. Le immagini associate alle emozioni erano prevalentemente femminili, creando un'ulteriore difficoltà per i partecipanti di sesso maschile nell'identificazione con gli stati emotivi proposti. Inoltre, è stato osservato che le persone che non consumavano caffè tendevano a rispondere alle domande mettendosi nei panni di chi ne faceva uso, introducendo un bias che ha complicato l'interpretazione dei risultati.

Infine, è emersa una difficoltà legata alla mancanza di conoscenza diffusa su quali alimenti o bevande contengano caffeina o teina. In futuro, fornire una lista dettagliata di questi elementi potrebbe migliorare la comprensione e l'accuratezza delle risposte da parte degli utenti.

Questi limiti evidenziano aree che necessitano di miglioramenti e suggeriscono possibili direzioni per ricerche future nel campo della robotica sociale.

5.2 Scenari futuri

Concluso l'esperimento con il robot Pepper, si è già avviato un progetto parallelo che prevede l'impiego del robot NAO¹⁶, con l'obiettivo di replicare l'esperimento originale. L'interesse principale risiede nella valutazione dell'influenza che l'uso di un robot con caratteristiche fisiche e abilità differenti possa esercitare sugli utenti. In particolare, si intende verificare se i risultati ottenuti con Pepper possano essere confermati o se l'interazione con NAO porti a cambiamenti significativi nelle risposte e nei comportamenti degli utenti.

NAO, sviluppato da SoftBank Robotics, si distingue per le sue caratteristiche fisiche compatte e umanoidi, con una statura di circa 58 cm, una mobilità articolata e una capacità espressiva visiva e motoria più marcata rispetto a Pepper. A differenza di Pepper, che si basa principalmente sulla comunicazione verbale e la presenza fisica, NAO ha una struttura che enfatizza movimenti dinamici e interattivi, grazie alle sue 25 articolazioni motorizzate e alla possibilità di eseguire gesti complessi e coordinati. Le sue abilità includono il riconoscimento vocale e visivo, la sintesi vocale e la capacità di eseguire gesti che rafforzano l'espressività emotiva durante le interazioni con gli utenti.

Questo esperimento con NAO prevede non solo la verifica della persuasività e della capacità empatica del robot, ma si pone anche l'obiettivo di esplorare come le differenze nelle caratteristiche fisiche e comportamentali del robot possano influenzare la percezione degli utenti. NAO, per esempio, utilizzerà un sistema gestuale per accompagnare l'espressione delle emozioni, integrando movimenti del corpo che rappresentano visivamente i suoi stati emotivi. Questi gesti aggiuntivi si prevede possano migliorare la comprensione emotiva da parte degli utenti, contribuendo a una maggiore immersione e coinvolgimento nella conversazione.

Un'altra implementazione che differenzierà l'esperimento con NAO riguarda l'uso del modulo *ALLeds*¹⁷ all'interno del codice preesistente. Questo modulo permette di variare il colore dei LED presenti sugli occhi del robot in base all'emozione espressa. Ad esempio, i LED potranno assumere tonalità calde per esprimere emozioni positive o tonalità fredde per emozioni negative. Questa aggiunta mira a offrire un ulteriore livello di feedback visivo, aiutando gli utenti a riconoscere con maggiore chiarezza l'emozione che NAO sta cercando di comunicare, migliorando così la qualità dell'interazione.

Sul fronte metodologico, è prevista una revisione del questionario somministrato agli utenti. A differenza dell'esperimento con Pepper, in cui i questionari pre e post-test erano ravvicinati, il nuovo progetto prevede che il questionario venga posto in due momenti distinti e distanti nel tempo. L'intento è quello di osservare se l'intervallo temporale permette di cogliere variazioni più evidenti nelle opinioni e negli atteggiamenti dei partecipanti, consentendo di misurare l'effetto della persuasione in modo più accurato e significativo.

Un'altra modifica rilevante riguarda la problematica emersa durante l'esperimento precedente, dove alcuni utenti manifestavano incertezze riguardo al contenuto di caffeina e teina nelle diverse bevande. Per colmare questa lacuna, sarà introdotta una nuova strategia di presentazione visiva. Durante l'interazione, verranno mostrate agli utenti immagini di varie bevande, tra cui ginseng, Coca Cola, tè e altre bibite che contengono stimolanti come caffeina e teina (vedi figura 14). Questo approccio visivo aiuterà a chiarire i dubbi e a fornire un contesto informativo più chiaro per le risposte degli utenti.

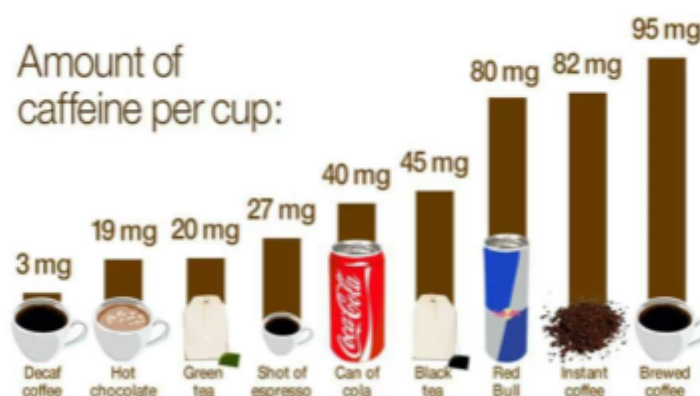


Figura 14: contenuto di caffeina presente nelle singole bevande. Da sinistra verso destra: caffè decaffeinato, cioccolata calda, tè verde, tazzina di caffè, lattina di coca-cola, tè nero, Red Bull, caffè istantaneo, tazza di caffè.

Fonte:

<https://www.ilcaffeespressoitaliano.com/2013/il-contenuto-di-caffeina-nelle-principali-bevande-energetiche/>

Nel complesso, il nuovo progetto con NAO non si limita a replicare l'esperimento condotto con Pepper, ma si arricchisce di ulteriori elementi di interazione, come gesti, feedback visivi attraverso i LED e un miglioramento del processo di raccolta dei dati. Queste innovazioni mirano a esplorare in modo più approfondito l'impatto della robotica sociale sugli esseri umani, mettendo in luce come diversi aspetti del design robotico possano influenzare la qualità e l'efficacia delle interazioni uomo-robot.

6. Riferimenti bibliografici e digitali

6.1 Bibliografia

Adam, M., Jibrin, A. G., & Ibrahim, D. (2019). Caffeine and Sleep Quality as Covariant of Academic Performance among Undergraduate Biology Education Students: Counselling Intervention. *International Journal of Scientific and Research Publications*, 9(3), 490-495.

10.29322/IJSRP.9.03.2019.p8780

Adda, S., Azouze, I., & Chaabane, F. (2022). Evaluation of Contextual Machine Translation: Case Study of Google Translate and Reverso Context Translate. *Ibn Khaldoun University of Tiaret*, 1-79.

Borenstein, J., & Arkin, R. C. (2017). Nudging for good: robots and the ethical appropriateness of nurturing empathy and charitable behavior. *AI & Society*, 32(4), 499-507.

Breazeal, C. (2000). *Sociable Machines: Expressive Social Exchange Between Humans and Robots* [Doctoral Dissertation. Department of Electrical Engineering and Computer Science.]. MIT. George M. Sprowls Award for Outstanding Ph.D. Thesis in Computer Science.

Breazeal, C. (2003). Emotion and sociable humanoid robots. *International Journal of Human-Computer Studies*, 59(1-2), 119-155. 10.1016/S1071-5819(03)00018-1

Breazeal, C. (2003). Towards social robots. *Robotics and Autonomous Systems*, 42(3-4), 167-175. <https://www.sciencedirect.com/science/article/abs/pii/S0921889002003731>

Cialdini, R. (2001). The Science of Persuasion. *Scientific American*, 284, 76-81.
10.1038/scientificamerican0201-76

Conti, D., Di Nuovo, S., Buono, S., & Di Nuovo, A. (2017). Robots in Education and Care of Children with Developmental Disabilities: A Study on Acceptance by Experienced and Future Professionals. *International Journal of Social Robotics*, 9(1), 51-62.

10.1007/s12369-016-0359-6

Dautenhahn, K. (2007). Socially intelligent robots: dimensions of human-robot interaction. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 362(1480), 679-704. <https://doi.org/10.1098/rstb.2006.2004>

de Carolis, B., D'errico, F., & Rossano, V. (2021). Pepper as a Storyteller: Exploring the Effect of Human vs. Robot Voice on Children's Emotional Experience. . 18th IFIP Conference on Human-Computer Interaction (INTERACT), 471-480. 10.1007/978-3-030-85616-8_27

Ekman, P., & Keltner, D. (1970). Universal facial expressions of emotion. *California mental health research digest*, 8(4), 151-158.

Fiore, F. (2019). Persuasione: cos'è e quali sono le principali tecniche usate per persuadere. *State of Mind*. <https://www.stateofmind.it/2019/02/persuasione-tecniche/>

Fiorini, L., D'Onofrio, G., Sorrentino, A., Loizzo, F. G.C., Russo, S., Giuliani, F., Sancarolo, D., & Cavallo, F. (2024). The Role of Coherent Robot Behavior and Embodiment in Emotion Perception and Recognition During Human-Robot Interaction: Experimental Study. *JMIR human factors*, 11(e45494). 10.2196/45494

Frijns, H. A., Schürer, O., & koeszegi, S. T. (2023). Communication Models in Human–Robot Interaction: An Asymmetric MODEL of ALterity in Human–Robot Interaction (AMODAL-HRI). *Int J of Soc Robotics*, 15, 473-500. 10.1007/s12369-021-00785-7

Gomes, A. A., Tavares, J., & de Azavedo, M. H. P. (2011). Sleep and Academic Performance in Undergraduates: A Multi-measure, Multi-predictor Approach. *Chronobiology International*, 28(9), 786-801. 10.3109/07420528.2011.606518

Goodrich, M., & Shultz, A. (n.d.). Human-Robot Interaction: A Survey. *Foundation and trends in Human-Computer Interaction*, Vol. 1, No. 3.

Il principio della riprova sociale secondo Robert Cialdini. (n.d.). Web Crew.
<https://webcrew.it/riprova-sociale-robert-cialdini/>

Introducing Whisper. (2022, September 21). OpenAI. Retrieved July 31, 2024, from
<https://openai.com/index/whisper/>

Jiao, W., Wang, W., Huang, J., Wang, X., & Tu, Z. (2023). Is ChatGPT A Good Translator? A Preliminary Study. *arXiv*.

Jong, M. d., Zhang, K., Roth, A. M., Rhodes, T., Schumacker, R., Zhou, C., Ferreira, S., Catucho, J., & Veloso, M. (2018). Towards a Robust Interactive and Learning Social Robot. In *Proceedings of the 17th International Conference on Autonomous Agents and MultiAgent Systems (AAMAS '18)*. International Foundation for Autonomous Agents and Multiagent Systems, 883-891.

Kirby, R., Forlizzi, J., & Simmons, R. (2010). Affective social robots. *Robotics and Autonomous Systems*, 58(3), 322-332. 10.1016/j.robot.2009.09.015

La ruota delle Emozioni di Plutchik. (2020). Six Seconds Italia.
<https://italia.6seconds.org/2020/10/la-ruota-delle-emozioni-di-plutchik/>

Lee, K. M., Peng, W., Jin, S.-A., & Yan, C. (2006). Can Robots Manifest Personality?: An Empirical Test of Personality Recognition, Social Responses, and Social Presence in Human–Robot Interaction. *Journal of Communication*, 56(4), 754-772.

10.1111/j.1460-2466.2006.00318.x

Li, H., Graesser, A. C., & Cai, Z. (2014). Comparison of Google Translation with Human Translation. The Florida AI Research Society.

Lieto, A., & Pozzato, G. L. (2019). A description logic framework for commonsense conceptual combination integrating typicality, probabilities and cognitive heuristics. *Journal of Experimental & Theoretical Artificial Intelligence*, 32(4), 1-36.

10.1080/0952813X.2019.1672799

Lieto, A., Pozzato, G. L., Striani, M., Zoia, S., & Damiano, R. (2023). DEGARI 2.0: A diversity-seeking, explainable, and affective art recommender for social inclusion. *Cognitive Systems Research*, 77(C), 1-17. 10.1016/j.cogsys.2022.10.001

Malinowska, J. K. (2021). What Does It Mean to Empathise with a Robot? *Minds & Machines*, 31, 361-376. 10.1007/s11023-021-09558-7

Mohammad, S. M., & Turney, P. D. (2013). Crowdsourcing a word-emotion association lexicon. *Computational Intelligence*, 29(3), 436-465. 10.48550/arXiv.1308.6297

Paiva, A., Leite, I., Boukricha, H., & Wachsmuth, I. (2017). Empathy in Virtual Agents and Robots: A Survey. *ACM Transactions Interactive Intelligent Systems (TiiS)*, 7(3), 1-40.

10.1145/2912150

Pandey, A. K., & Gelin, R. (2018). A Mass-Produced Sociable Humanoid Robot: Pepper: The First Machine of Its Kind. *IEEE Robotics & Automation Magazine*, 25(3), 40-48.

10.1109/MRA.2018.2833157

Paradedda, R. B., Martinho, C., & Paiva, A. (2020). Persuasion Strategies Using a Social Robot in an Interactive Storytelling Scenario. In *Proceedings of the 8th International Conference on Human-Agent Interaction*. Association for Computing Machinery, New York, NY, USA, 69-77.

Picard, R. W. (1997). *Affective computing*. MIT Press.

Plutchik, R. (1980). A general psychoevolutionary theory of emotion. *Theories of emotion*, Elsevier, 3-33.

Polakova, P., & Klimova, B. (2023). Using DeepL translator in learning English as an applied foreign language – An empirical pilot study. *Heliyon*, 9(8). 10.1016/j.heliyon.2023.e18595

Polillo, R. (2010). Che cos'è un prototipo. Retrieved August 20, 2024, from <http://www.rpolillo.it/faciledausare/Cap.9.htm>

Quattro, P. (2004). Testing agreement among multiple retars. *ResearchGate*, 145-151.

10.6092/issn.1973-2201/28

Simkus, J. (2024). Between-Subjects Design: Overview & Examples. *SimplyPsychology*.

Siragusa, G. (2020). *PROMPT ENGINEERING*. 1-57.

Sisini, V. (2024). Prompt engineering: come formulare richieste efficaci per avere risposte precise dalle IA. *Cultura e società digitali, Agenda Digitale*.

Suardiaz-Muro, M., Oretaga-Moreno, M., Morante-Ruiz, M., Monroy, M., Ruiz, M. A., Martín-Plasencia, P., & Vela-Bueno, A. (2023). Sleep quality and sleep deprivation: relationship with academic performance in university students during examination period. *Sleep and Biological Rhythms*, Springer Link, 21, 377-383. 10.1007/s41105-023-00457-1

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). Attention Is All You Need. 1-15. 10.48550/arXiv.1706.03762

Winkle, K., Lemaingnan, S., Caleb-Solly, P., Leonards, U., Turton, A., & Bremner, P. (2019). Effective Persuasion Strategies for Socially Assistive Robots. 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI), 277-285. 10.1109/HRI.2019.8673313.

Zafari, S., Schwaniger, I., Hirschmanner, M., Schmidbauer, C., Weiss, A., & Koeszegi, S. T. (2019). You Are Doing so Great!” – The Effect of a Robot’s Interaction Style on Self-Efficacy in HRI. 2019 28th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN), 1-7. 10.1109/RO-MAN46459.2019.8956437

6.2 Sitografia

1. *SoftBank Robotics*. (n.d.). Group Corp. Retrieved July 26, 2024, from <https://www.softbankrobotics.com/>
2. Google Traduttore. (n.d.). Google Translate. Retrieved July 31, 2024, from <https://translate.google.com/?hl=it&sl=auto&tl=it&op=translate>
3. *Traduzione in contesto - arabo, tedesco, inglese, spagnolo, francese, ebraico, giapponese, olandese, polacco, portoghese, rumeno, russo, svedese, turco, ucraino, italiano, coreano, cinese*. (n.d.). Reverso Context. Retrieved July 31, 2024, from <https://context.reverso.net/traduzione/>
4. *ChatGPT*. (n.d.). OpenAI. Retrieved July 31, 2024, from <https://openai.com/chatgpt/>
5. *Il miglior traduttore al mondo*. (n.d.). DeepL Translate. Retrieved July 31, 2024, from <https://www.deepl.com/it/translator>
6. *Introducing Whisper*. (2022, September 21). OpenAI. Retrieved July 31, 2024, from <https://openai.com/index/whisper/>
7. *JSON: cosa è, come si usa e con quali tool - Introduzione di base*. (2023, June 22). DigitalWorld. Retrieved July 31, 2024, from <https://www.digitalworlditalia.it/sviluppo-software/web/json-cosa-e-122893>
8. Kotlin Programming Language. Retrieved July 31, 2024, from <https://kotlinlang.org/>
9. Pepper SDK for Android — QiSDK. Retrieved July 31, 2024, from <https://qisdk.softbankrobotics.com/sdk/doc/pepper-sdk/index.html>
10. *Playground AI - Free AI Image Generator*. (n.d.). Retrieved August 20, 2024, from <https://playground.com/>
11. *Canva: Visual Suite for Everyone*. (n.d.). Retrieved August 20, 2024, from <https://www.canva.com/>

12. *Manuale operativo di design | 4.4. Personas: creare i profili degli utenti-tipo*. (n.d.). Docs Italia. Retrieved July 20, 2024, from <https://docs.italia.it/italia/designers-italia/manuale-operativo-design-docs/it/versione-corrente/doc/design-research/personas.html>
13. *Excel*. (n.d.). Microsoft. Retrieved July 25, 2024, from <https://www.microsoft.com/it-it/microsoft-365/p/excel/cfq7ttc0k7dx>
14. *Omogeneità della varianza valutare le uguali varianze in ANOVA a due vie*. (2024, June 11). FasterCapital. Retrieved July 27, 2024, from <https://fastercapital.com/it/contenuto/Omogeneita-della-varianza--valutare-le-uguali-varianze-in-ANOVA-a-due-vie.html>
15. *Understanding Autonomous Life settings — Aldebaran 2.4.3.28-r2 documentation*. (n.d.). SoftBank Robotics Documentation. Retrieved July 27, 2024, from http://doc.aldebaran.com/2-4/family/pepper_user_guide/life_pep.html
16. *NAO le robot humanoïde et programmable*. (n.d.). Aldebaran. Retrieved July 25, 2024, from <https://www.aldebaran.com/it/nao>
17. *ALLeds* —. (n.d.). NAO Software 1.14.5 documentation. Retrieved August 25, 2024, from <http://doc.aldebaran.com/1-14/naoqi/sensors/alleds.html>