

Empatía Percibida en Interacciones Humano-Agente: Comparación de Instrumentos y Desarrollo del Cuestionario CEPRS

Perceived Empathy in Human-Agent Interactions: Comparison of Instruments and Development of the CEPRS Questionnaire

Alberto Hitos-García

Dpto. Lenguajes y Sistemas
Informáticos
Universidad de Granada
Granada, España
albertohitos@ugr.es

Francisco Luis Gutiérrez-Vela

Dpto. Lenguajes y Sistemas
Informáticos
Universidad de Granada
Granada, España
fgutierr@ugr.es

Patricia Paderewski-Rodríguez

Dpto. Lenguajes y Sistemas
Informáticos
Universidad de Granada
Granada, España
patricia@ugr.es

André Satoshi Kawamoto

Dpto. Acadêmico de
Computação
Universidade Tecnológica
Federal de Paraná
Campo Mourão-PR, Brasil
kawamoto@utfpr.edu.br

Recibido: 12.11.2025 | Aceptado: 2.12.2025

DOI: <https://doi.org/10.65234/interaccion.132>

Palabras Clave

Empatía Percibida
Interacción Persona-
Ordenador (IPO)
Instrumentos de Medición
Análisis Comparativo

Resumen

La empatía percibida en las interacciones con agentes computacionales es un tema emergente y de gran importancia en la Interacción Persona-Ordenador (IPO). Este estudio presenta un análisis comparativo de los instrumentos existentes para medir la empatía percibida en agentes computacionales y, a partir de las limitaciones detectadas, propone un nuevo cuestionario adaptado a la interacción con adultos mayores. Al evaluar los fundamentos teóricos, la adaptabilidad y los procesos de validación de las herramientas revisadas, el artículo identifica fortalezas y debilidades clave, especialmente en relación con la evaluación de señales multimodales y la naturaleza dinámica de la interacción. Además, se destaca la necesidad de instrumentos accesibles y sensibles al contexto emocional y comunicativo de distintos grupos de usuarios, especialmente los adultos mayores. Los hallazgos respaldan el desarrollo de enfoques de medición más inclusivos y específicos, con el fin de fomentar interacciones Humano-Agente más efectivas, significativas y confiables.

Keywords

Perceived Empathy
Human-Computer Interaction
(HCI)
Assessment Instruments
Comparative Analysis

Abstract

Perceived empathy in interactions with computational agents is a significant and emerging topic in Human-Computer Interaction (HCI). This study provides a comparative analysis of existing instruments for evaluating perceived empathy in computational agents and based on the identified limitations, introduces a new questionnaire specifically adapted for interactions with older adults. By examining each instrument's theoretical foundations, adaptability, and validation processes, the paper identifies key strengths and weaknesses, particularly regarding multimodal and dynamic interaction contexts. Furthermore, it highlights the need for accessible and context-sensitive assessment tools tailored to diverse user groups, especially older adults. The findings support the development of more inclusive, domain-relevant approaches to empathy assessment, fostering the design of more meaningful, trustworthy, and engaging Human-Agent Interactions.

1. Introducción

En los últimos años, la presencia de agentes computacionales en la vida diaria ha aumentado significativamente. Los agentes basados en Inteligencia Artificial (IA) son habituales

en diversos sectores: en la atención médica, interactúan con pacientes ofreciendo apoyo emocional y asistiendo en decisiones médicas; en la educación, sirven como asistentes de enseñanza y monitorizan el progreso de los estudiantes; y

en las tareas cotidianas, agentes como Alexa se han vuelto omnipresentes (Lugrin et al., 2022).

Aunque estos agentes están diseñados para asistir e interactuar con los usuarios, la forma en que los usuarios perciben estas interacciones juega un papel importante en la satisfacción, confianza y efectividad del usuario (Birmingham et al., 2022; Yalçın et al., 2024; Y. Yang et al., 2017).

La empatía (entender y compartir las emociones de otra persona) es vital para el éxito de las interacciones sociales. Para los agentes computacionales, es esencial que los usuarios perciban empatía para construir confianza, promover el compromiso y asegurar experiencias positivas (Charrier et al., 2019; Chaves & Gerosa, 2021; Cuadra et al., 2024; Paiva et al., 2017; B. Yang et al., 2023).

Sin embargo, es necesaria una adaptación de la empatía al contexto específico de las Interacciones Humano-Agente, ya que los agentes computacionales, al ser sistemas artificiales, no pueden experimentar emociones como lo hacen los humanos (Birmingham et al., 2022; Cuadra et al., 2024; Paiva et al., 2017; Park & Whang, 2022). En este contexto, la empatía es una simulación diseñada para mejorar la experiencia del usuario y hacer la interacción más natural y satisfactoria (Birmingham et al., 2022; Cuadra et al., 2024; James et al., 2018; Paiva et al., 2017; Park & Whang, 2022).

Sin embargo, esta simulación de empatía plantea preguntas sobre la autenticidad de las respuestas del agente y el riesgo de crear expectativas poco realistas al usuario. Si un agente expresa empatía a través del lenguaje, el tono de voz o las expresiones faciales, los usuarios podrían interpretar estas expresiones como evidencia de sentimientos genuinos, lo que lleva a una comprensión distorsionada de las capacidades del agente (Concannon & Tomalin, 2024; Cuadra et al., 2024).

Desde un punto de vista práctico, los desarrolladores se enfrentan al desafío de equilibrar respuestas empáticas y realistas con la necesidad de gestionar las expectativas del usuario. Los agentes deben reconocer y responder apropiadamente a las emociones de los usuarios, al tiempo que evitan crear la ilusión de que el agente posee una comprensión emocional profunda (Birmingham et al., 2022; Concannon & Tomalin, 2024; Roshanaei et al., 2024).

Para mitigar estos desafíos, los usuarios deben ser conscientes de que están interactuando con un sistema artificial que simula la empatía y no con un humano capaz de experimentar emociones genuinas. Educar a los usuarios sobre las limitaciones de la empatía computacional puede ayudar a

gestionar las expectativas y prevenir malentendidos (Birmingham et al., 2022; Cuadra et al., 2024).

En la actualidad, existe una variedad de herramientas para medir la empatía en las interacciones humanas, como la Escala de Empatía de Hogan (*Hogan Empathy Scale*, HES) (Hogan, 1969), el Índice de Reactividad Interpersonal (*Interpersonal Reactivity Index*, IRI) (Davis & others, 1980), la Escala de Empatía del Terapeuta (*Therapist Empathy Scale*, TES) (Decker et al., 2014) y la Escala de Empatía del Médico de Jefferson (*Jefferson Scale of Physician Empathy*, JSPE) (Hojat et al., 2018).

La mayoría de estas herramientas son cuestionarios basados en afirmaciones que permiten a los participantes u observadores independientes evaluar subjetivamente una interacción basada en la conversación.

Si bien instrumentos como el IRI se utilizan ampliamente en las interacciones humanas, su aplicación directa a agentes artificiales es limitada. Además, a pesar de que las métricas de empatía humana están bien establecidas en la psicología, la psiquiatría y la neurociencia, la traducción directa de estas métricas al dominio de la IA no es trivial y requiere adaptaciones cuidadosas considerando las diferencias fundamentales entre humanos y agentes computacionales (Concannon & Tomalin, 2024; Yalçın et al., 2024).

Por tanto, a pesar de su reconocida importancia, la investigación sobre la empatía en sistemas interactivos aún necesita instrumentos de medición estandarizados y validados que capturen la complejidad de la empatía humana y se adapten a las particularidades de la IPO (Concannon & Tomalin, 2024; Putta et al., 2022; Roshanaei et al., 2024; Schmidmaier et al., 2024; Yalçın et al., 2024). Esta falta de instrumentos fiables dificulta la comparación de diferentes sistemas y el desarrollo de otros nuevos que incorporen la empatía en sus características.

Dada la complejidad de medir la empatía en las interacciones entre humanos y agentes computacionales, este artículo tiene como objetivo ofrecer una revisión y comparación detallada entre los diferentes instrumentos existentes para medir la empatía percibida en agentes computacionales, destacando los desafíos de adaptar las herramientas tradicionales a interacciones no humanas y enfatizando cómo abordar las lagunas actuales, explorando oportunidades futuras para expandir esta área de investigación.

Adicionalmente, a partir del estudio comparativo, se propone un nuevo cuestionario específicamente diseñado para evaluar la empatía percibida por adultos mayores en su interacción

con agentes interactivos sociales. Esta propuesta surge de la necesidad de contar con instrumentos que no solo consideren las particularidades de la empatía en Interacciones Humano-Agente, sino también las características cognitivas, emocionales y comunicativas propias de esta población (Kawamoto et al., 2014).

Con ello, este trabajo pretende responder a la siguiente pregunta de investigación: ¿Qué limitaciones presentan los instrumentos actuales de medición de empatía en agentes artificiales y cómo deben adaptarse para evaluar eficazmente la interacción con adultos mayores?

2. Búsqueda de instrumentos de evaluación de la empatía y criterios de comparación

Se realizó una búsqueda bibliográfica temática para identificar instrumentos de evaluación de la empatía percibida en las Interacciones Humano-Agente. Siguiendo los principios de (Braun & Clarke, 2006), este método permitió una exploración iterativa y flexible, priorizando la relevancia conceptual y la diversidad metodológica sobre la exhaustividad estadística de otras metodologías de búsqueda.

La búsqueda se realizó en las bases de datos IEEE Xplore, ACM Digital Library, PubMed, SpringerLink y arXiv, acotando los resultados a la última década, para garantizar su vigencia tecnológica. Se seleccionaron palabras clave propias de los temas identificados y se combinaron utilizando operadores booleanos para equilibrar la amplitud y la especificidad en los resultados de la búsqueda. La búsqueda incluyó términos como "*empathy measurement*", "*perceived empathy assessment tools*" o "*adapted empathy scales for computational agents*", entre otros. El proceso de refinamiento no fue lineal, sino guiado por temas esenciales identificados preliminarmente.

Este enfoque temático priorizó la inclusión, permitiendo examinar herramientas de diversos contextos y modalidades. Se incluyeron artículos y preprints basados en su relevancia y su enfoque en instrumentos adaptados para la Interacción Humano-Agente. Por el contrario, se excluyeron los estudios estrictamente limitados a la empatía de humano a humano (o interpersonal) o que carecían de suficientes detalles metodológicos.

Esta metodología destacó la riqueza teórica y la diversidad de los instrumentos de evaluación de la empatía, al tiempo que identificó lagunas que justifican una exploración futura.

La muestra final de instrumentos se sometió a un análisis comparativo estructurado en base a nueve criterios definidos

para evaluar su idoneidad en contextos de Interacción Humano-Agente y que proporcionan una base para que los posibles investigadores interesados en adoptar un instrumento puedan seleccionar el más adecuado a su objetivo:

1. **Adaptación de una Métrica Existente:** verifica si el instrumento es una adaptación de métricas encontradas en la literatura. Este criterio es esencial para comprender la base teórica del instrumento y su validez previa, ya que los instrumentos adaptados pueden tener un historial de uso y validación, lo que contribuye a su fiabilidad y aplicabilidad.
2. **Tipo de Sistema Evaluado:** identifica el tipo de sistema interactivo computacional (diálogo, interactivo, robot social u otro) al que se dirige el instrumento. Los diferentes sistemas pueden influir en la percepción de la empatía, y la comprensión de estas particularidades ayuda a contextualizar las interacciones y adaptar los instrumentos para diversas tecnologías.
3. **Dimensiones/Ítems:** evalúa las dimensiones e ítems medidos por el instrumento, así como el número de ítems en cada dimensión. Este criterio garantiza un análisis exhaustivo de la empatía percibida, comprendiendo qué aspectos específicos de la empatía aborda el instrumento. El número de dimensiones e ítems también indica si el instrumento se aplica a audiencias específicas, como los adultos mayores, que pueden sufrir una carga cognitiva en cuestionarios extensos.
4. **Punto de Vista:** el punto de vista adoptado para medir la empatía puede ser desde el usuario, desde un observador externo o desde un experto. El punto de vista puede influir en los resultados y la interpretación de los datos, proporcionando información variada sobre la interacción y la empatía percibida.
5. **Tipo de Cuestionario:** como por ejemplo la escala Likert o las preguntas exploratorias. El tipo de cuestionario influye en cómo se recopilan y analizan los datos, con cuestionarios estructurados en escalas conocidas que facilitan la comparación y la interpretación de los resultados.
6. **Validación:** verifica si el instrumento se sometió a una validación empírica, como el análisis factorial, de fiabilidad (Alfa de Cronbach u Omega de McDonald) y de validez convergente/divergente. La validación empírica es fundamental para garantizar la fiabilidad y la solidez científica del instrumento.

7. **Limitaciones:** identifica posibles restricciones en el método de aplicación, el tipo de sistema evaluado o el tipo de interacción estudiada. Comprender las limitaciones del instrumento es esencial para identificar áreas de mejora y restricciones.
8. **Enfoque en el Lenguaje:** analiza si el instrumento se centra solo en el lenguaje o considera otras señales (como expresiones faciales y gestos). Este criterio es significativo para situaciones en las que solo se utilizan agentes conversacionales con una fuerte dependencia de la comunicación verbal (hablada o escrita), como los *chatbots* que utilizan el Procesamiento del Lenguaje Natural.
9. **Modalidades de Interacción:** que evalúa las modalidades utilizadas por el instrumento (texto, voz, multimodal). Considerar las modalidades de interacción ayuda a comprender la flexibilidad y la aplicabilidad del instrumento en diferentes contextos de uso, ya que las diferentes modalidades pueden afectar de forma única la percepción de la empatía.

Los criterios se organizaron en tres grupos temáticos para resumir la comparación y facilitar la presentación y el análisis comparativo de los instrumentos. Esta comparación puede consultarse en los anexos, en forma de tablas.

La primera tabla presenta los criterios generales relacionados con las características contextuales de los instrumentos, como el tipo de sistema evaluado, las modalidades de interacción que admite, el año de publicación y si adapta o no métricas preexistentes en la interacción interpersonal.

La segunda tabla incluye información sobre la estructura y el contenido de los instrumentos, incluido el número de dimensiones e ítems, el enfoque en el lenguaje u otras señales, y la perspectiva adoptada (usuario, observador externo o experto).

Finalmente, la tercera tabla describe los aspectos metodológicos como el tipo de cuestionario utilizado, los procesos de validación empírica llevados a cabo y las limitaciones identificadas.

Esta división tiene como objetivo organizar la información de manera clara y objetiva, optimizando la comprensión y la comparación entre los instrumentos.

La siguiente sección proporciona una descripción completa de estos instrumentos, destacando sus características únicas, dimensiones y posibles contribuciones a la evaluación de la empatía en sistemas computacionales.

3. Herramientas para medir la empatía

3.1. Modelo Basado en QEAE para Evaluar la Empatía en las Interacciones Humano-Robot

Tisseron *et al.* tuvieron como objetivo explorar y validar un modelo de empatía aplicado a las interacciones entre humanos y robots, utilizando un Cuestionario de Empatía y Auto-empatía (*Questionnaire of Empathy and Auto-Empathy*, QEAE) (Tisseron et al., 2015). Su principal propuesta es comprender cómo se percibe la empatía en las interacciones con robots y avatares, contribuyendo a la aceptación y efectividad de los robots de asistencia, particularmente en contextos de tratamiento y rehabilitación.

Según los autores, el QEAE mide las diferentes dimensiones de la empatía durante las interacciones y puede utilizarse en dos contextos experimentales: un estudio psicológico con una población no clínica y una iniciativa de investigación interdisciplinaria centrada en la interacción entre adultos mayores y robots humanoides.

Los resultados esperados de este proyecto incluyeron una comprensión más profunda de las relaciones empáticas que los humanos pueden desarrollar con los robots, lo que puede tener implicaciones significativas para el diseño y la implementación de robots de asistencia. La investigación buscó validar el QEAE e identificar qué aspectos de las interacciones y características del robot influyen en la experiencia empática humana.

El QEAE se desarrolló como una herramienta adaptable para medir la empatía en las interacciones con robots y avatares, lo que indica una adaptación métrica para diferentes contextos de interacción. Al abordar tanto robots como avatares, este modelo proporciona información sobre los matices de la empatía en diferentes tipos de entidades artificiales, haciéndolo versátil y completo. Esta adaptabilidad es clave, ya que permite que el instrumento sea aplicable en varios escenarios que involucren interacciones Humano-Robot o Humano-Avatar.

El QEAE consta de cuatro dimensiones: **auto-empatía**, **empatía directa**, **empatía recíproca** y **empatía intersubjetiva**. Cada dimensión contiene cuatro componentes (acción, emoción, cognición y asistencia) que suman un total de 16 ítems. Esta estructura detallada asegura la captura de múltiples facetas de la empatía, proporcionando un marco robusto para la evaluación.

El QEAE está diseñado para captar la perspectiva del usuario, lo que permite una evaluación directa de los procesos

emocionales y cognitivos involucrados en estas interacciones, mejorando la relevancia de los hallazgos.

Los autores no especifican explícitamente la escala adoptada por el instrumento, pero señalan que el QEAЕ aún está en proceso de validación, con estudios experimentales planificados para confirmar su efectividad y relevancia en las interacciones Humano-Robot. Esta validación en curso refuerza la credibilidad científica del instrumento y garantiza su fiabilidad en diversos contextos.

El instrumento se aplica a interacciones multimodales, incluyendo texto, voz y otras señales no verbales. Esta aplicabilidad multimodal refleja cómo los humanos interactúan con robots y avatares, evaluando de manera integral la empatía en estos contextos.

Los autores admiten que el QEAЕ aún está en proceso de validación y reconocen algunas limitaciones. Este proceso puede dificultar la generalización de los resultados hasta que se complete la validación. Además, las interacciones con robots pueden presentar desafíos únicos que el modelo podría no abordar completamente, destacando áreas para un futuro refinamiento.

3.2. RoPE (Robot's Perceived Empathy)

Charrier *et al.* desarrollaron la escala de Empatía Percibida del Robot (*Robot's Perceived Empathy*, RoPE) para medir las percepciones humanas de la empatía en las interacciones con robots, abordando una brecha en la investigación de la empatía robótica (Charrier et al., 2019). Una ventaja clave de la escala RoPE es su estandarización, que evita la aplicación directa de métricas de empatía humana al reconocer las diferencias en las capacidades emocionales entre humanos y robots.

La escala adapta las métricas utilizadas para medir la empatía humana al contexto de la interacción Humano-Robot, evitando formulaciones que se basen en habilidades de las que los robots carecen.

La escala RoPE consta de dos dimensiones: **Comprensión Empática y Respuesta Empática**, con 18 ítems distribuidos en estas dimensiones y cuatro ítems extra. Esta estructura permite una evaluación integral de la empatía, capturando tanto los componentes cognitivos como los emocionales a través de un formato de cuestionario tipo Likert. Este formato es ideal para obtener respuestas matizadas y proporciona un enfoque estandarizado para medir la empatía en diferentes estudios y contextos.

Expertos en ciencia cognitiva y robótica han llevado a cabo una validación preliminar. Este proceso de validación en curso es importante para establecer la credibilidad científica de la escala y garantizar su efectividad en diversas aplicaciones. Los autores afirman que hay planes para probar la fiabilidad y validez de la versión francesa de la escala en un experimento futuro, utilizando un robot Cozmo para interacciones grabadas que varían entre reacciones empáticas y neutrales.

De cierta manera, la necesidad de una mayor validación para garantizar la fiabilidad y sensibilidad de la escala representa una limitación. Además, la aplicación de la escala puede estar influenciada por sesgos relacionados con las capacidades emocionales de los robots, lo que debe considerarse al interpretar los resultados.

Aunque el artículo original de la escala RoPE no proporciona criterios de evaluación detallados, (Daher et al., 2020) utilizaron RoPE para medir la empatía percibida en las interacciones con un *chatbot* de asistencia médica. Sus resultados demostraron la efectividad de la escala para capturar diferencias en la percepción de la empatía basadas en variaciones en las respuestas del *chatbot*.

El estudio concluyó que RoPE es una herramienta válida y valiosa para evaluar la interacción Humano-*Chatbot*, particularmente en escenarios de atención médica. Destacó la sensibilidad de la escala a los ajustes en las respuestas del *chatbot* con respecto a la empatía percibida. Sin embargo, los autores recomendaron estudios futuros en diversos contextos con muestras más grandes para validar estos hallazgos y abordar posibles limitaciones y la generalización de la herramienta.

3.3. Modelo de Evaluación de la Empatía en Sistemas Autónomos

El modelo de evaluación de la empatía propuesto por Urakami *et al.* proporciona un marco integral para comprender cómo los usuarios perciben las expresiones empáticas en las interacciones con sistemas autónomos (Urakami et al., 2019). El modelo consta de ocho dimensiones que capturan aspectos clave de la empatía, lo que permite un análisis detallado de las respuestas de los usuarios: **expresión de sentimientos propios** (la capacidad del sistema para comunicar sus emociones), **expresión de saber lo que siente el otro** (reconocer y articular las emociones del usuario), **ayuda** (la voluntad del sistema de ofrecer asistencia), **mostrar interés** (demostrar atención y preocupación por las necesidades del usuario), **tomar la perspectiva del otro** (comprender la situación del usuario desde su punto de vista), **mostrar consideración** (reconocer y respetar las emociones del usuario), **comprensión situacional** (empatía cognitiva,

evaluar la situación emocional del usuario) y **acuerdo** (alinearse con las emociones u opiniones del usuario).

La empatía se evalúa mediante una encuesta con 72 ítems. Aunque los autores no especifican el número de ítems por dimensión, la estructura integral permite una evaluación detallada, capturando aspectos clave de la Interacción Persona-Ordenador.

Esta herramienta adapta las definiciones e instrumentos existentes para medir la empatía, centrándose en elementos relevantes para las interacciones con sistemas autónomos, asegurando su idoneidad para evaluar la empatía en diálogos. La evaluación se lleva a cabo desde la perspectiva del usuario. En un estudio realizado por los autores, los participantes evaluaron las expresiones empáticas de un sistema autónomo basándose en sus experiencias, asegurando que los hallazgos reflejen interacciones reales de los usuarios.

La validación del modelo implicó la comparación de las evaluaciones de cuatro investigadores, los cuales llegaron a las mismas conclusiones. Los resultados mostraron que las dimensiones de empatía cognitiva y asistencia, como "ayuda" y "mostrar interés", fueron calificadas más positivamente que las dimensiones afectivas como "expresión de sentimientos propios". Esto sugiere que los usuarios valoran la comprensión y el apoyo práctico por encima de las expresiones emocionales en las interacciones con sistemas autónomos.

El cuestionario utiliza una escala tipo Likert, donde los participantes califican las afirmaciones de 0 a 10, indicando su grado de acuerdo. Este formato captura la intensidad de la respuesta y proporciona un método estandarizado para medir la empatía.

El instrumento se centra en el lenguaje, específicamente el texto, ya que el estudio se realizó a través de una encuesta *online* que evaluaba afirmaciones escritas. No se mencionan las señales no verbales, lo que sugiere un área potencial de expansión para incluir señales de comunicación multimodal, con el instrumento diseñado principalmente para diálogos textuales.

La consistencia interna se confirmó mediante el alfa de Cronbach, arrojando valores satisfactorios, lo que indica dimensiones fiables y consistentes. Además, se utilizó ANOVA para comparar las evaluaciones entre dimensiones, revelando diferencias significativas en las percepciones de los participantes.

Este modelo mejora la comprensión de las interacciones entre humanos y sistemas autónomos y proporciona un marco

valioso para la investigación de la empatía en IA. Las dimensiones identificadas y la metodología de evaluación pueden guiar el desarrollo de sistemas más efectivos y empáticos, mejorando potencialmente la aceptación y efectividad del usuario en aplicaciones del mundo real.

Los autores reconocen limitaciones, como que los participantes evalúan la empatía mediante escenarios hipotéticos, lo que puede no reflejar interacciones reales, y la falta de exploración de las diferencias individuales en las respuestas a las expresiones empáticas.

3.4. Social Service Robot Interaction Trust (SSRIT)

La Confianza en la Interacción con Robots de Servicio Social (*Social Service Robot Interaction Trust*, SSRIT) mide la confianza de los consumidores en las interacciones con robots sociales de inteligencia artificial en contextos de prestación de servicios (Chi et al., 2021). Aunque desarrollada con un enfoque en la confianza, la escala presenta dimensiones directamente relacionadas con la percepción de la empatía.

La confianza a menudo se reconoce como una condición facilitadora de la empatía, permitiendo a los usuarios involucrarse más profundamente en las interacciones y atribuir intenciones positivas al agente social. Por lo tanto, el SSRIT contribuye indirectamente a comprender cómo se percibe la empatía en los escenarios de Interacción Humano-Robot.

La estructura multidimensional del SSRIT captura características clave para analizar la empatía en las interacciones computacionales. La dimensión "**propensión a confiar**" incluye factores como la familiaridad y la autoeficacia, que influyen en la voluntad de los usuarios de interpretar los comportamientos del agente como empáticos. Por ejemplo, las personas con mayor familiaridad tecnológica tienden a reducir las barreras cognitivas, mejorando su recepción de comportamientos empáticos.

La dimensión "**función y diseño de robot confiable**" incluye atributos como el antropomorfismo y el rendimiento, cruciales para demostrar empatía, ya que las características similares a las humanas y la funcionalidad constante son claves para evaluar las respuestas emocionales simuladas. Finalmente, la dimensión "**tarea y contexto de servicio confiable**" considera factores situacionales como el riesgo percibido y el ajuste entre el robot y el servicio, que influyen en la receptividad emocional y empática de los usuarios.

Además, la solidez metodológica del SSRIT, que combina indicadores formativos y reflexivos, ofrece una perspectiva

integral que puede facilitar futuras adaptaciones o expansiones de escalas centradas en la empatía percibida. En este sentido, la integración de SSRIT en el alcance de este estudio permite una comprensión teórica y operativa más profunda de la relación entre la confianza y la empatía en las interacciones Humano-Robot.

3.5. Uso de Instrumentos Existentes para Medir la Empatía Humana

La utilización del TEQ en su forma original, desarrollado para interacciones humanas, subraya una limitación significativa en su aplicación a entornos virtuales (Kroes et al., 2022).

Kroes *et al.* exploraron la empatía hacia los agentes virtuales, centrándose en la influencia de las historias de personificación y las características individuales del usuario. Utilizaron el Cuestionario de Empatía de Toronto (*Toronto Empathy Questionnaire*, TEQ), una escala Likert que mide la empatía como un constructo multidimensional, enfatizando los componentes afectivos y emocionales. Los participantes también completaron una encuesta posterior al experimento evaluando sus reacciones emocionales al agente virtual basándose en los componentes del TEQ.

El TEQ es una métrica no adaptada diseñada originalmente para interacciones humanas y no se ha modificado específicamente para evaluar la empatía hacia los agentes virtuales. Su uso en su forma original destaca una limitación significativa para la aplicación en entornos virtuales.

El instrumento incluye seis dimensiones: **Contagio Emocional** (*EmCon*), **Comprensión Emocional** (*EmUnd*), **Sensibilidad** (*Sens*), **Activación Fisiológica Simpática** (*SympPhy*), **Altruismo** (*AltEmp*) y **Comportamiento Empático de Orden Superior**. Aunque conceptualizadas inicialmente para interacciones de Humano-Humano, estas dimensiones proporcionan un marco integral para evaluar diversas facetas de la empatía. El instrumento utiliza una escala Likert de 5 puntos para capturar la intensidad de los sentimientos y actitudes de los participantes hacia el agente virtual.

El estudio evaluó un entorno inmersivo de realidad virtual (RV) donde los participantes observaron a un agente virtual mostrando emociones. Este entorno inmersivo es clave para comprender la empatía en un entorno más realista, proporcionando información sobre la dinámica de la Interacción Persona-Ordenador.

La interacción ocurrió en un entorno multimodal, con participantes observando al agente expresar emociones en RV. A pesar de la falta de interacción directa (por ejemplo,

diálogo o voz), la configuración permitió examinar la empatía en entornos virtuales.

El estudio concluyó que la empatía hacia los agentes virtuales está influenciada por las características individuales, pero la personificación por sí sola puede no aumentar la empatía. Los autores enfatizan la necesidad de considerar las diferencias individuales en el desarrollo de Interacciones Humano-Agente efectivas, sugiriendo que se requiere más investigación para adaptar las medidas de empatía como el TEQ a contextos de agentes computacionales.

El TEQ se utilizó desde la perspectiva del usuario, capturando percepciones subjetivas de empatía basadas en sus experiencias y sentimientos hacia el agente. Este enfoque centrado en el usuario es vital para evaluar la efectividad de los agentes virtuales en la evocación de respuestas empáticas. El TEQ incorpora componentes emocionales y conductuales, midiendo la empatía en un contexto más amplio. Sin embargo, es posible que no tenga completamente en cuenta las señales de comunicación no verbal, que son clave en las interacciones cara a cara.

Si bien el TEQ está validado en interacciones humanas, su aplicabilidad a agentes virtuales sigue sin abordarse, lo que presenta una notable laguna. Las limitaciones incluyen su diseño original para interacciones humanas, el tamaño de muestra pequeño y los posibles sesgos de los participantes que se centran en la apariencia del agente en lugar de en sus emociones.

3.6. Marco Multidimensional para la Empatía en Diálogos Interactivos

(Xu & Jiang, 2024) propusieron un marco de evaluación de la empatía diseñado para capturar la complejidad de la empatía en los diálogos interactivos, enfatizando su naturaleza colaborativa entre el hablante y el oyente. El marco consta de dos dimensiones principales: **Empatía Expresada** y **Empatía Percibida**.

La Empatía Expresada se refiere a las intenciones comunicativas del hablante, evaluadas mediante la identificación de actos de habla que transmiten empatía.

La Empatía Percibida adapta las definiciones psicológicas recientes a los diálogos orientados a tareas, evaluando la empatía desde la perspectiva del oyente. Esta dimensión incluye cuatro aspectos: **Compromiso**, que mide la percepción del oyente de la participación del hablante; **Comprensión**, que evalúa qué tan bien siente el oyente que se entienden sus emociones y su situación; **Simpática**, que evalúa la visión del oyente de la empatía y la respuesta

adecuada del hablante; y **Utilidad**, que evalúa si la comunicación del hablante aborda los problemas centrales de la conversación.

Estos aspectos interconectados forman un marco integral que refleja la naturaleza multifacética de la empatía en las interacciones sociales. La aplicación de este marco a los diálogos de servicio al cliente reveló correlaciones significativas entre la empatía percibida y la satisfacción del diálogo.

Este enfoque mejora la comprensión de la empatía en la comunicación y ofrece un método robusto para evaluar automáticamente la empatía, contribuyendo al desarrollo de sistemas de diálogo más efectivos y satisfactorios.

El instrumento adapta las métricas existentes para medir la empatía en los diálogos, considerando tanto las intenciones expresadas como la empatía percibida. Es muy adecuado para analizar las interacciones de servicio al cliente, lo que lo hace muy relevante en este contexto.

El marco fue validado utilizando un conjunto de datos interno de 2.000 diálogos de servicio al cliente anotados. Evaluadores humanos identificaron 16 intenciones comunicativas expresadas y cuatro aspectos de la empatía percibida, calificándolos en una escala Likert-5. El análisis mostró que la empatía expresada y percibida están interconectadas, y la empatía percibida influye directamente en la satisfacción de la conversación.

El estudio también encontró que los clasificadores basados en modelos de lenguaje ajustados por instrucciones superaron a los métodos de *prompting* y a otros enfoques competitivos. Estos resultados demuestran la efectividad del marco para medir la empatía en los diálogos, respaldando su aplicación en contextos de comunicación del mundo real.

La empatía percibida se evaluó desde la perspectiva del oyente, mientras que las intenciones expresadas se evaluaron desde la perspectiva del hablante.

Si bien el instrumento se centra en las características lingüísticas, también tiene en cuenta la dinámica social entre los participantes del diálogo, reconociendo la empatía como un proceso colaborativo. La evaluación se basa en texto, específicamente en diálogos de servicio al cliente. Sin embargo, el enfoque es adaptable a otras modalidades, ofreciendo un marco flexible para evaluar la empatía en diversos contextos de comunicación. Esta visión más amplia garantiza la inclusión de múltiples dimensiones de la empatía, extendiéndose más allá de la comunicación verbal.

Como limitaciones, los autores destacan los desafíos en la medición de la empatía debido a su naturaleza implícita, los estilos de expresión que pueden ser específicos del dominio y la sutil distinción entre la empatía expresada y percibida. Además, la efectividad del modelo puede variar según el contexto de interacción, lo que indica áreas que requieren un mayor refinamiento.

3.7. Heurísticas de Diseño Empático y Emocional en Tecnologías Sanitarias

Las heurísticas de diseño empático y emocional desarrolladas por (Borycki et al., 2024) ofrecen un conjunto de pautas destinadas a mejorar las interacciones usuario-sistema, particularmente dentro de las tecnologías sanitarias. Estas heurísticas se derivaron de una revisión exhaustiva de la literatura que identificó las mejores prácticas y principios que conducen a la creación de experiencias de usuario más positivas y emocionalmente satisfactorias.

El proceso de desarrollo implicó un análisis cualitativo de los datos recopilados, seguido de una revisión y validación por parte de un panel de expertos en información sanitaria y factores humanos, asegurando que las heurísticas sean tanto basadas en evidencia como contextualmente relevantes.

Las heurísticas se organizan en categorías que abordan elementos críticos del diseño empático, incluida la **Personalización**, la **Mensajería**, la **Participación** y la **Usabilidad**. Por ejemplo, se enfatiza la personalización como un factor clave para mejorar la percepción de empatía de los usuarios al adaptar la presentación de la información a las necesidades individuales. Además, se recomienda el uso de comunicación sencilla, junto con imágenes y visualizaciones, para aumentar la participación y la satisfacción del usuario. Estas pautas forman el diseño de la interfaz y sirven como herramientas de evaluación para medir la efectividad de las interacciones en términos de empatía y participación del usuario.

Estas heurísticas se aplican a diversas modalidades de interacción, como interfaces gráficas, textuales y multimodales, así como a robots sociales y asistentes virtuales. Su versatilidad les permite ser utilizadas en diversos contextos, que van desde aplicaciones de atención médica hasta plataformas de *e-learning*, donde la empatía y la experiencia del usuario juegan un papel crítico para las interacciones exitosas. Si bien el artículo no menciona una validación empírica formal de las heurísticas, el proceso de desarrollo, basado en una revisión de la literatura y la validación de expertos, proporciona una base sólida para su

uso práctico y efectividad en la promoción de interacciones empáticas.

El instrumento está diseñado para sistemas interactivos e interfaces de usuario, particularmente en tecnologías sanitarias. Si bien no se limita a un tipo de sistema específico, la herramienta es aplicable en contextos que involucran diálogos e interacciones con el usuario, lo que la hace adecuada para una amplia gama de aplicaciones. Esta adaptabilidad es vital para evaluar interfaces que tienen como objetivo fomentar interacciones más empáticas y emocionalmente atractivas.

Desde el punto de vista de la evaluación, el instrumento se desarrolla desde la perspectiva del usuario, centrándose en cómo las interacciones son percibidas y experimentadas por los usuarios finales. Además, los observadores externos y los expertos pueden participar en la evaluación durante el proceso de validación, mejorando el análisis y ofreciendo una comprensión más completa de la efectividad de las heurísticas.

El enfoque del instrumento se extiende más allá del lenguaje para incluir una variedad de señales y componentes de diseño, como mensajes, imágenes y participación, que trascienden la comunicación verbal.

Aunque la validación empírica no se aborda explícitamente en el documento, un panel de expertos revisó y validó las heurísticas, lo que indica cierto grado de validación a pesar de la falta de un proceso empírico formal en estudios de campo. Si bien el documento no analiza directamente las limitaciones, es razonable suponer que puede haber desafíos para generalizar las heurísticas a diferentes contextos de interacción. Podría ser necesaria una mayor validación en diversos escenarios. Además, la efectividad de las heurísticas puede depender del tipo de sistema y la complejidad de las interacciones, lo que debe considerarse al aplicarlas en contextos futuros.

3.8. Empathy Scale for Human-Computer Communication (ESHCC)

La Escala de Empatía para la Comunicación Humano-Ordenador (*Empathy Scale for Human-Computer Communication*, ESHCC) (Concannon & Tomalin, 2024) representa una adaptación de una métrica existente, la Escala de Empatía del Terapeuta (*Therapist Empathy Scale*, TES), desarrollada originalmente para medir la empatía en las interacciones Humano-Humano, especialmente en contextos terapéuticos. Esta adaptación es importante para evaluar la empatía en los sistemas de diálogo (SD), como los asistentes virtuales y los *chatbots*.

El marco ESHCC considera los aspectos matizados que definen la empatía en los contextos digitales, incluyendo: **Relevancia**, la capacidad del sistema para responder adecuadamente a las emociones y necesidades del usuario, asegurando que la interacción siga siendo significativa; **Fluidez**, la naturalidad y la suavidad de las respuestas del sistema, incluso si ocasionalmente sacrifica la expresividad emocional en aras de la coherencia; **Conciencia Emocional**, la capacidad del sistema para reconocer y responder apropiadamente a las emociones expresadas por el usuario, lo que refleja su sensibilidad a las señales emocionales; **Conexión Interpersonal**, la sensación de cercanía y comprensión que el usuario percibe durante la interacción, que es clave para fomentar la buena relación.

La evaluación se realiza desde la perspectiva de un observador externo, lo que permite un análisis imparcial de la empatía demostrada por el sistema a lo largo del diálogo. Para este propósito, el ESHCC utiliza una escala tipo Likert, donde los evaluadores califican cada ítem en una escala del 1 al 7, proporcionando una medida cuantitativa de la empatía percibida.

Los autores validaron la herramienta a través de un estudio empírico que implicó la aplicación del instrumento en diversas interacciones con sistemas de diálogo. Los resultados indicaron que el ESHCC posee propiedades psicométricas sólidas, incluida una alta fiabilidad y validez convergente. El análisis factorial confirmatorio reveló que las dimensiones del ESHCC se agrupan coherentemente, lo que respalda la estructura teórica propuesta.

Según los autores, este instrumento puede servir como una herramienta valiosa para desarrolladores e investigadores que buscan mejorar la Interacción Persona-Ordenador, permitiendo la identificación de áreas de mejora en las respuestas del sistema.

Como limitaciones, el ESHCC puede capturar solo parcialmente la empatía en las interacciones basadas en texto. Enfatiza el lenguaje como medio principal de evaluación, centrándose en las características léxicas, textuales y sintácticas en lugar de las señales no verbales o paralingüísticas que son más relevantes en las interacciones cara a cara.

El instrumento se aplica tanto a interacciones textuales como a transcripciones de interacciones de voz, aunque el énfasis principal está en la comunicación textual. Este enfoque permite un análisis más profundo de la empatía en contextos

de comunicación mediados por tecnología, lo que contribuye a la comprensión de la interacción Humano-Máquina.

3.9. Perceived Empathy of Technology Scale (PETS)

La Escala de Empatía Percibida de la Tecnología (*Perceived Empathy Technology Scale*, PETS) (Schmidmaier et al., 2024) mide la empatía percibida de los sistemas interactivos, con el objetivo de cerrar la brecha en la evaluación de la empatía tecnológica. PETS se desarrolló a través de un proceso estructurado que incluyó contribuciones de expertos y pruebas con usuarios, asegurando su relevancia para los sistemas interactivos.

El instrumento consta de 10 ítems, divididos en dos dimensiones principales: **Sensibilidad Emocional** (PETS-ER) y **Comprensión y Confianza** (PETS-UT), cada una de las cuales contiene cinco ítems que evalúan diferentes facetas de la empatía percibida. La recopilación de datos se facilita a través de una escala Likert, lo que respalda el análisis estadístico.

PETS se creó utilizando un enfoque de abajo hacia arriba (*bottom-up*), incorporando comentarios de expertos y pruebas con usuarios, lo que lo hace altamente específico para los sistemas interactivos. Se aplica a tecnologías que demuestran empatía, como asistentes virtuales y robots sociales, en una variedad de aplicaciones. Los participantes califican la empatía del sistema desde su propia perspectiva en varios escenarios.

La validación empírica de PETS se llevó a cabo a través de análisis factoriales exploratorios y confirmatorios, asegurando la solidez y fiabilidad del instrumento. El marco considera tanto el lenguaje como la respuesta emocional y la comprensión del usuario, abarcando aspectos tanto afectivos como cognitivos. PETS es versátil, aplicable a interacciones de texto, voz y multimodales, lo que amplía su aplicabilidad a diversos escenarios de interacción.

A pesar de la validación estadística, los autores reconocen la complejidad de la empatía en las interacciones humanas y la posible variación en la interpretación de las dimensiones de la empatía en diferentes contextos.

4. Desafíos, limitaciones e implicaciones éticas

4.1. Limitaciones de los Instrumentos Existentes y Adaptados

Si bien ha habido avances en el desarrollo y la adaptación de herramientas de medición de la empatía para agentes computacionales, persisten varias limitaciones. Muchos

instrumentos adaptados todavía arrastran suposiciones de las interacciones interpersonales (Humano-Humano), que pueden no alinearse completamente con la dinámica de las relaciones Humano-Agente, lo que lleva a posibles sesgos.

Por ejemplo, las nociones preconcebidas de los usuarios sobre la tecnología pueden influir en su percepción de la empatía, lo que a menudo resulta en una proyección poco realista de las capacidades emocionales en el agente. Además, las capacidades emocionales y cognitivas limitadas de los sistemas de IA actuales restringen la precisión de las evaluaciones de empatía, ya que estos agentes solo pueden simular emociones en lugar de experimentarlas. Estos desafíos complican la medición y la interpretación de la empatía percibida, a menudo comprometiendo la validez de los resultados en diversos contextos.

4.2. Implicaciones Éticas y Sociales

La simulación de la empatía por parte de los sistemas de IA introduce importantes preocupaciones éticas. ¿Es ético crear sistemas que parecen empatizar sin experimentar realmente emociones? Si bien estas simulaciones pueden ofrecer beneficios potenciales, también podrían inducir a error a los usuarios al hacerles pensar que el agente tiene una comprensión similar a la humana, lo que lleva a expectativas poco realistas. Esto es particularmente problemático en áreas sensibles como la atención médica y el apoyo a la salud mental, donde los usuarios podrían formar apegos emocionales o depender del agente para obtener apoyo emocional. Estas dinámicas exigen una cuidadosa consideración de la transparencia en el diseño de los sistemas de IA, asegurando que los usuarios sean conscientes de las limitaciones de la empatía de la IA. Además, el potencial de que la IA empática se utilice indebidamente con fines manipuladores, como influir en el comportamiento del consumidor, subraya la necesidad de directrices éticas con respecto a la simulación de la empatía en los sistemas de IA.

4.3. Lagunas y Oportunidades de Investigación

Los instrumentos de empatía revisados en este estudio tratan predominantemente la empatía como una característica estática evaluada después de la interacción. Si bien tales enfoques proporcionan información valiosa, a menudo no tienen en cuenta la naturaleza dinámica y evolutiva de la empatía durante las Interacciones Humano-Agente en tiempo real.

A pesar de que herramientas como el QEAE y el RoPE exhiben metodologías estructuradas, sus diseños se centran en escenarios predefinidos, lo que limita su capacidad para capturar la sensibilidad emocional momento a momento, particularmente crítica en dominios como la atención médica

o el servicio al cliente. La evaluación en tiempo real permitiría una comprensión más profunda de cómo se desarrolla la empatía en las diferentes etapas de la interacción.

Además, si bien a menudo se discute la generalización en diversos dominios, los hallazgos de instrumentos como el ESHCC sugieren que la precisión dentro de dominios específicos tiene mayor importancia que la aplicabilidad universal. Por ejemplo, los sistemas de atención médica pueden requerir herramientas altamente especializadas para abordar los matices emocionales de las interacciones paciente-médico, mientras que las tecnologías educativas pueden beneficiarse de medidas adaptadas al apoyo motivacional. Los instrumentos deben centrarse en la adaptabilidad dentro de contextos definidos, asegurando su sensibilidad a las demandas únicas de cada dominio.

Otra oportunidad clave radica en la integración de la variabilidad cultural e individual en el diseño de instrumentos. Los estudios sobre marcos como el SSRIT demuestran que la familiaridad de los usuarios con la tecnología y los antecedentes culturales influyen significativamente en la percepción de la empatía. La incorporación de tales factores en las herramientas de evaluación podría mejorar su inclusividad y relevancia, particularmente para las aplicaciones globales. Además, las dimensiones de la empatía, como la expresividad emocional frente a la comprensión cognitiva, pueden variar en importancia dependiendo del contexto cultural del usuario, lo que subraya la necesidad de diseños flexibles que se adapten a diversas interpretaciones.

La investigación futura puede desarrollar instrumentos capaces de medir la empatía de forma dinámica al tiempo que garantiza la precisión específica del dominio y la adaptabilidad cultural. Estos avances respaldarán la creación de sistemas Humano-Agente que sean empáticos en el diseño y sensibles a la naturaleza multifacética de las interacciones con los usuarios, lo que contribuirá a mejorar la confianza, la satisfacción y el compromiso.

5. Propuesta del cuestionario CEPRS para adultos mayores

La revisión comparativa realizada en las secciones anteriores pone de manifiesto que, si bien los instrumentos existentes ofrecen contribuciones valiosas para la medición de la empatía percibida, presentan limitaciones importantes cuando se aplican en contextos que involucran a adultos mayores y robots sociales.

En particular, se observó que varios cuestionarios adaptan constructos de la empatía humana sin considerar las particularidades de la interacción con agentes

computacionales, lo que puede derivar en interpretaciones sesgadas sobre la capacidad emocional del agente. Este fenómeno es especialmente relevante en usuarios con escasa experiencia en el uso de la tecnología, como suele ser el caso de los adultos mayores.

Además, algunos instrumentos presentan una carga cognitiva elevada (por ejemplo, escalas con hasta 72 ítems o formulaciones abstractas), lo cual dificulta su aplicación al colectivo de adultos mayores, donde es imprescindible evitar la sobrecarga cognitiva.

Por último, se identificó una falta de herramientas que integren explícitamente dimensiones asociadas a la compañía social y la sensación de seguridad, aspectos especialmente relevantes en la interacción con adultos mayores.

La mayoría de los instrumentos revisados se centran en validar las competencias emocionales o cognitivas del agente, pero descuidan la evaluación del vínculo afectivo y la fiabilidad. Estas dimensiones resultan determinantes en la interacción con adultos mayores, un colectivo donde la aceptación tecnológica no depende tanto de la complejidad del diálogo, sino de la capacidad del robot para mitigar la sensación de soledad (compañía) y ofrecer un entorno de interacción predecible y libre de riesgos (seguridad).

En respuesta a estas limitaciones, se propone el **Cuestionario de Empatía Percibida en Robots Sociales (CEPRS)**, diseñado para evaluar la experiencia empática desde la perspectiva del adulto mayor en interacciones con robots sociales. Esta propuesta no solo selecciona las dimensiones más relevantes de la literatura revisada, sino que las adapta a las características cognitivas y emocionales del adulto mayor (Kawamoto et al., 2014), priorizando la claridad conceptual y la simplicidad en la redacción de las afirmaciones sobre la abstracción técnica.

Este cuestionario se basa en 4 dimensiones extraídas del estudio comparativo:

1. **Resonancia emocional** (Concannon & Tomalin, 2024; Schmidmaier et al., 2024; Tisseron et al., 2015; Urakami et al., 2019): Evalúa en qué medida el usuario percibe que el robot expresa emociones de forma coherente y adecuada durante la interacción, generando una sensación de cercanía afectiva. Incluye la percepción de naturalidad, calidez y credibilidad en las expresiones emocionales del agente.
2. **Comprensión empática** (Charrier et al., 2019; Concannon & Tomalin, 2024; Schmidmaier et al., 2024; Urakami et al., 2019; Xu & Jiang, 2024): Mide la

capacidad atribuida al robot para entender las necesidades, intenciones y mensajes del usuario, ofreciendo respuestas apropiadas y ajustadas al contexto de la conversación.

- 3. **Confianza y seguridad** (Chi et al., 2021; Schmidmaier et al., 2024): Evalúa el grado de seguridad y fiabilidad que el robot transmite durante la interacción, tanto en la consistencia de sus respuestas como en sus movimientos (si corresponde).
- 4. **Socialización y compañía** (Chi et al., 2021; Tisseron et al., 2015; Xu & Jiang, 2024): Mide la percepción del usuario sobre el robot como agente social que facilita el vínculo, el acompañamiento y la interacción significativa. Considera el disfrute de la conversación, la sensación de presencia social y el potencial del robot para fomentar la conexión con otras personas.

Para garantizar su adecuación metodológica al grupo objetivo, el CEPRS reduce drásticamente la carga cognitiva respecto a propuestas anteriores. Para ello, su diseño prioriza ítems breves, formulados en lenguaje claro, sin tecnicismos y con ejemplos integrados cuando es relevante. La estructura final, compuesta por cuatro ítems por dimensión, se detalla en la Tabla 1.

Se utiliza una escala Likert de cinco puntos, complementada con un pequeño bloque de preguntas de sí o no que permiten matizar aspectos expresivos y contextuales del robot.

Esta combinación reduce la carga cognitiva, facilita la aplicación en contextos naturales (por ejemplo, sesiones grupales o talleres en residencias) y favorece respuestas más fiables.

Asimismo, su diseño permite ser aplicado tras una única sesión de interacción o a lo largo de experiencias repetidas, lo que lo convierte en una herramienta flexible para estudios longitudinales sobre aceptación, adaptación y relación con agentes sociales.

El cuestionario completo puede consultarse en los anexos.

Tabla 1. Dimensiones e ítems del cuestionario CEPRS

ID	Resonancia emocional
RE1	El robot me pareció cercano y amable.
RE2	Pensé que el robot se preocupaba por cómo me siento.
RE3	Las emociones del robot me parecieron creíbles y naturales.
RE4	La cara del robot mostraba emociones apropiadas con lo que decía (por ejemplo, si decía algo triste no tuviera una cara alegre).
ID	Comprensión empática
CE1	El robot entendió bien lo que le pedía.
CE2	Sentí que el robot prestaba atención a lo que le decía.
CE3	Las respuestas del robot tenían sentido para mí.
CE4	El robot me ayudó cuando lo necesité.
ID	Confianza y seguridad
CS1	Sentí que podía confiar en el robot.
CS2	Me sentí seguro hablando con el robot.
CS3	Sentí que el robot se preocupaba por la seguridad de los demás, especialmente al moverse.
CS4	Me sentí tranquilo respecto a cómo se comportaba el robot.
ID	Socialización y compañía
SC1	Disfruté de la conversación con el robot.
SC2	El robot me ayudó a socializar con otras personas.
SC3	Sentí que el robot era un buen compañero.
SC4	Me sentí acompañado mientras hablaba con el robot.

6. Conclusiones y trabajo futuro

La empatía es un elemento fundamental en el diseño de sistemas de IA que sean efectivos, atractivos y capaces de fomentar la confianza y la aceptación del usuario. La medición precisa de la empatía juega un papel fundamental en la creación de interacciones más parecidas a las humanas, especialmente a medida que la IA se integra cada vez más en dominios sensibles como la atención médica, la educación y la asistencia social. Esta revisión es un recurso valioso para aquellos que buscan seleccionar herramientas de evaluación de la empatía adecuadas para su trabajo.

Este artículo ha examinado el panorama de los instrumentos para medir la empatía percibida en la IPO, abordando sus fundamentos teóricos, marcos metodológicos y aplicaciones prácticas.

A través de un análisis comparativo de las herramientas existentes y adaptadas, se han destacado sus fortalezas, relevancia específica del dominio y limitaciones. Si bien es

exhaustiva, esta revisión temática reconoce la posible influencia de los sesgos de selección subjetivos, que pueden afectar la reproducibilidad de sus hallazgos.

Este estudio enfatiza la necesidad de desarrollar herramientas de medición de la empatía sensibles al contexto para abordar mejor los desafíos que plantean los agentes computacionales. Sin embargo, durante el estudio se ha observado una falta de herramientas adaptadas a la interacción con adultos mayores, un contexto de gran relevancia para los agentes computacionales, ya que estos pueden contribuir a afrontar desafíos sociales cada vez más urgentes, como el envejecimiento de la población, la soledad no deseada y la necesidad de apoyar la autonomía en ámbitos de salud y vida diaria.

Como respuesta a esta carencia, este trabajo propone el Cuestionario de Empatía Percibida en Robots Sociales (CEPRS), que ofrece un enfoque específico para evaluar la interacción de robots sociales con adultos mayores, integrando dimensiones críticas como la compañía y la seguridad que habían sido olvidadas en la literatura previa.

En trabajos futuros, se llevará a cabo la validación empírica del cuestionario CEPRS con adultos mayores en contextos de interacción real con robots en entornos sociales.

Esta validación incluirá un análisis de fiabilidad interna (por ejemplo, el coeficiente alfa de Cronbach) y un análisis factorial confirmatorio (CFA) para evaluar la coherencia estructural de las dimensiones propuestas.

Además, se realizarán estudios longitudinales que permitan observar cómo evolucionan las percepciones de empatía a medida que aumenta la familiaridad del usuario con el agente.

Agradecimientos

Este trabajo se ha llevado a cabo en el contexto del proyecto PID2022-1367790B-C33 (PLEISAR-SOCIAL) con fondos del Ministerio de Ciencia, Innovación y Universidades (MICIU/AEI/10.13039/501100011033) y FEDER/EU y del proyecto C-ING-179-UGR23 financiado por la Consejería de Universidades, Investigación e Innovación y por el programa FEDER Andalucía 2021-2027.

Referencias

- Birmingham, C., Perez, A., & Mataric, M. (2022). Perceptions of Cognitive and Affective Empathetic Statements by Socially Assistive Robots. *ACM/IEEE International Conference on Human-Robot Interaction, 2022-March*, 323–331. <https://doi.org/10.1109/HRI53351.2022.9889386>
- Borycki, E. M., Kletke, R., le Nobel, C., McWilliams, G., Whitehouse, S., & Kushniruk, A. W. (2024). Empathetic and Emotive Design Heuristics. In *pHealth 2024* (pp. 80–84). IOS Press. <https://doi.org/10.3233/SHTI240062>
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>
- Charrier, L., Rieger, A., Galdeano, A., Cordier, A., Lefort, M., & Hassas, S. (2019). The RoPE Scale: a Measure of How Empathic a Robot is Perceived. *Proceedings of the 14th ACM/IEEE International Conference on Human -Robot Interaction*, 656–657. <https://doi.org/10.1109/HRI.2019.8673082>
- Chaves, A. P., & Gerosa, M. A. (2021). How Should My Chatbot Interact? A Survey on Social Characteristics in Human-Chatbot Interaction Design. *International Journal of Human-Computer Interaction*, 37(8), 729–758. <https://doi.org/10.1080/10447318.2020.1841438>
- Chi, O. H., Jia, S., Li, Y., & Gursoy, D. (2021). Developing a formative scale to measure consumers' trust toward interaction with artificially intelligent (AI) social robots in service delivery. *Computers in Human Behavior*, 118, 106700. <https://doi.org/10.1016/j.chb.2021.106700>
- Concannon, S., & Tomalin, M. (2024). Measuring perceived empathy in dialogue systems. *AI & SOCIETY*, 39(5), 2233–2247. <https://doi.org/10.1007/s00146-023-01715-z>
- Cuadra, A., Wang, M., Stein, L. A., Jung, M. F., Dell, N., Estrin, D., & Landay, J. A. (2024). The Illusion of Empathy? Notes on Displays of Emotion in Human-Computer Interaction. *Proceedings of the CHI Conference on Human Factors in Computing Systems*, 1–18. <https://doi.org/10.1145/3613904.3642336>
- Daher, K., Casas, J., Khaled, O. A., & Mugellini, E. (2020). Empathic Chatbot Response for Medical Assistance. *Proceedings of the 20th ACM International Conference on Intelligent Virtual Agents*, 1–3. <https://doi.org/10.1145/3383652.3423864>
- Davis, M. H., & others. (1980). A multidimensional approach to individual differences in empathy. *JSAS Catalog of Selected Documents in Psychology*, 10, 85. https://www.uv.es/friasnav/Davis_1980.pdf
- Decker, S. E., Nich, C., Carroll, K. M., & Martino, S. (2014). Development of the therapist empathy scale. *Behavioural and Cognitive Psychotherapy*, 42(3), 339–354. <https://doi.org/10.1017/S1352465813000039>
- Hogan, R. (1969). Development of an empathy scale. *Journal of Consulting and Clinical Psychology*, 33(3), 307–316. <https://doi.org/10.1037/h0027580>
- Hojat, M., DeSantis, J., Shannon, S. C., Mortensen, L. H., Speicher, M. R., Bragan, L., LaNoe, M., & Calabrese, L. H. (2018). The Jefferson Scale of Empathy: a nationwide study of measurement properties, underlying components, latent variable structure, and national norms in medical students. *Advances in Health Sciences Education*, 23, 899–920. <https://doi.org/10.1007/s10459-018-9839-9>

- James, J., Watson, C. I., & MacDonald, B. (2018). Artificial Empathy in Social Robots: An analysis of Emotions in Speech. *Proceedings of the 27th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, 632–637. <https://doi.org/10.1109/ROMAN.2018.8525652>
- Kroes, K., Saccardi, I., & Masthoff, J. (2022). Empathizing with virtual agents: the effect of personification and general empathic tendencies. *Proceedings of the IEEE International Conference on Artificial Intelligence and Virtual Reality*, 73–81. <https://doi.org/10.1109/AIVR56993.2022.00017>
- Lugrin, B., Pelachaud, C., & Traum, D. (Eds.). (2022). *The Handbook on Socially Interactive Agents: 20 years of Research on Embodied Conversational Agents, Intelligent Virtual Agents, and Social Robotics Volume 2: Interactivity, Platforms, Application* (1st ed., Vol. 48). Association for Computing Machinery.
- Paiva, A., Leite, I., Boukricha, H., & Wachsmuth, I. (2017). Empathy in Virtual Agents and Robots: A Survey. *ACM Trans. Interact. Intell. Syst.*, 7(3), 11:1–11:40. <https://doi.org/10.1145/2912150>
- Park, S., & Whang, M. (2022). Empathy in Human–Robot Interaction: Designing for Social Robots. *International Journal of Environmental Research and Public Health*, 19(3), 1889. <https://doi.org/10.3390/ijerph19031889>
- Putta, H., Daher, K., Kamali, M. El, Khaled, O. A., Lalanne, D., & Mugellini, E. (2022). Empathy scale adaptation for artificial agents: a review with a new subscale proposal. *Proceedings of the 8th International Conference on Control, Decision and Information Technologies*, 1, 699–704. <https://doi.org/10.1109/CoDIT55151.2022.9803993>
- Roshanaei, M., Rezapour, R., & El-Nasr, M. S. (2024). *Talk, Listen, Connect: Navigating Empathy in Human-AI Interactions*. <https://doi.org/10.48550/arXiv.2409.15550>
- Schmidmaier, M., Rupp, J., Cvetanova, D., & Mayer, S. (2024). Perceived Empathy of Technology Scale (PETS): Measuring Empathy of Systems Toward the User. *Proceedings of the CHI Conference on Human Factors in Computing Systems*, 1–18. <https://doi.org/10.1145/3613904.3642035>
- Tisseron, S., Tordo, F., & Baddoura, R. (2015). Testing Empathy with Robots: A Model in Four Dimensions and Sixteen Items. *International Journal of Social Robotics*, 7(1), 97–102. <https://doi.org/10.1007/s12369-014-0268-5>
- Urakami, J., Moore, B. A., Sutthithatip, S., & Park, S. (2019). Users' Perception of Empathic Expressions by an Advanced Intelligent System. *Proceedings of the 7th International Conference on Human-Agent Interaction*, 11–18. <https://doi.org/10.1145/3349537.3351895>
- Xu, Z., & Jiang, J. (2024). *Multi-dimensional Evaluation of Empathetic Dialog Responses*. <https://doi.org/10.48550/arXiv.2402.11409>
- Yalçın, Ö. N., Utz, V., & DiPaola, S. (2024). Empathy through Aesthetics: Using AI Stylization for Visual Anonymization of Interview Videos. *Proceedings of the 3rd Empathy-Centric Design Workshop: Scrutinizing Empathy Beyond the Individual*, 63–68. <https://doi.org/10.1145/3661790.3661803>
- Yang, B., Sun, Y., & Shen, X.-L. (2023). Understanding AI-based customer service resistance: A perspective of defective AI features and tri-dimensional distrusting beliefs. *Information Processing & Management*, 60(3), 103257. <https://doi.org/10.1016/j.ipm.2022.103257>
- Yang, Y., Ma, X., & Fung, P. (2017). Perceived Emotional Intelligence in Virtual Agents. *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, 2255–2262. <https://doi.org/10.1145/3027063.3053163>

Anexos

Anexo 1. Comparación de Instrumentos

Tabla 2. Características Generales

Instrumento	Año de Publicación	¿Adapta una Métrica Existente?	Tipo de Sistema Evaluado	Modalidades de Interacción
Tisseron <i>et al.</i> (QEAE)	2015	No se indica explícitamente	Robots Sociales y Avatares	Multimodal
Charrier <i>et al.</i> (RoPE)	2019	Adapta escalas de empatía utilizadas en ciencia cognitiva	Robots Sociales	Multimodal
Urakami <i>et al.</i> (Empathy Assessment Model)	2019	Ítems adaptados de marcos de empatía cognitiva	Sistemas interactivos	Texto
Chi <i>et al.</i> (SSRIT)	2021	No	Robots Sociales	Multimodal
Kroes <i>et al.</i> (TEQ)	2022	No (utiliza TEQ sin adaptación)	Sistemas de Realidad Virtual	Multimodal
Concannon y Tomalin (ESHCC)	2023	Adaptación de TES	Sistemas de Diálogo	Texto
Borycki <i>et al.</i> (Design Heuristics)	2024	No	Sistemas de Salud	Multimodal
Schmidmaier <i>et al.</i> (PETS)	2024	No	Sistemas Interactivos	Multimodal
Xu y Jiang	2024	Adapta definiciones de empatía psicológica	Sistemas de Diálogo de Servicio al Cliente	Texto

Tabla 3. Estructura y Contenido

Instrumento	Dimensiones/Ítems	Enfoque en el Lenguaje	Punto de Vista del Evaluador
Tisseron <i>et al.</i> (QEAE)	4 dimensiones / 16 ítems	No	Perspectiva del Usuario
Charrier <i>et al.</i> (RoPE)	2 dimensiones / 18 ítems	No	Perspectiva del Usuario
Urakami <i>et al.</i> (Empathy Assessment Model)	8 dimensiones / 72 ítems	Sí	Perspectiva del Usuario
Chi <i>et al.</i> (SSRIT)	3 dimensiones / No indicado	No	Perspectiva del Usuario
Kroes <i>et al.</i> (TEQ)	6 dimensiones del TEQ	No	Perspectiva del Usuario
Concannon y Tomalin (ESHCC)	4 dimensiones / No indicado	Sí	Observador Externo
Borycki <i>et al.</i> (Design Heuristics)	No applicable	No	Perspectiva del Usuario y Expertos
Schmidmaier <i>et al.</i> (PETS)	2 dimensiones / 10 ítems	No	Perspectiva del Usuario
Xu y Jiang	2 dimensiones / 16 intenciones / 4 aspectos	Sí	Perspectiva Dual (Usuario y Externo)

Tabla 4. Características Metodológicas

Instrumento	Tipo de Cuestionario	Validación	Limitaciones
Tisseron <i>et al.</i> (QEAE)	No especificado	En proceso	Validación limitada, conjunto de ítems pequeño
Charrier <i>et al.</i> (RoPE)	Escala Likert	Preliminar	Limitado a robots, necesita pruebas más amplias
Urakami <i>et al.</i> (Empathy Assessment Model)	Escala Likert	Alfa de Cronbach, ANOVA	Escenarios hipotéticos
Chi <i>et al.</i> (SSRIT)	Reflexivo y Formativo	Validación empírica	Enfoque indirecto en la empatía
Kroes <i>et al.</i> (TEQ)	Escala Likert	Validación empírica	No adaptado para agentes
Concannon y Tomalin (ESHCC)	Escala Likert	Análisis Factorial Confirmatorio	Limitado a la interacción textual
Borycki <i>et al.</i> (Design Heuristics)	No aplicable	Revisión de expertos	Sin validación empírica formal
Schmidmaier <i>et al.</i> (PETS)	Escala Likert	Análisis Factorial Exploratorio y Confirmatorio	Aplicabilidad específica al contexto
Xu y Jiang	Escala Likert	Validación empírica en conjunto de datos anotado	Enfoque específico del dominio en diálogos

Anexo 2. Propuesta de Cuestionario

Cuestionario para evaluar la Empatía Percibida en Robot Sociales (CEPRS)

A continuación, encontrará una serie de afirmaciones. Por favor, lea cada una con atención y marque con una X el número que mejor refleje su grado de acuerdo con la afirmación.

Totalmente en desacuerdo	En desacuerdo	Ni de acuerdo ni en desacuerdo	De acuerdo	Totalmente de acuerdo
1	2	3	4	5

Afirmaciones	Grado de acuerdo				
	1	2	3	4	5
<i>El robot me pareció cercano y amable.</i>					
<i>Pensé que el robot se preocupaba por cómo me siento.</i>					
<i>Las emociones del robot me parecieron creíbles y naturales.</i>					
<i>El robot entendió bien lo que le pedía.</i>					
<i>Sentí que el robot prestaba atención a lo que le decía.</i>					
<i>Las respuestas del robot tenían sentido para mí.</i>					
<i>Sentí que podía confiar en el robot.</i>					
<i>Me sentí seguro hablando con el robot.</i>					
<i>Sentí que el robot se preocupaba por la seguridad de los demás, especialmente al moverse.</i>					
<i>Disfruté de la conversación con el robot.</i>					
<i>El robot me ayudó a socializar con otras personas.</i>					
<i>Sentí que el robot era un buen compañero.</i>					

Lea atentamente cada afirmación y marque con una X en función de si está de acuerdo o no. También puede elegir la opción NS/NC (No sabe / No contesta) en caso de no estar seguro o no querer responder.

Afirmaciones	Sí	No	NS/NC
<i>La cara del robot mostraba emociones apropiadas con lo que decía (por ejemplo, si decía algo triste no tuviera una cara alegre).</i>			
<i>El robot me ayudó cuando lo necesité.</i>			
<i>Me sentí tranquilo respecto a cómo se comportaba el robot.</i>			
<i>Me sentí acompañado mientras hablaba con el robot.</i>			