CUERPO B

UNIVERSIDAD DE PALERMO
FACULTAD DE DISEÑO Y COMUNICACIÓN
DOCTORADO EN DISEÑO

TESIS DOCTORAL

Diseño de interacción: análisis de la percepción de consciencia en chatbots inteligentes.

Una propuesta de pauta de evaluación fundamentada en la semiótica visual, el affordance y las heurísticas de interfaz

Autora:
Aracelli Muñoz Baldi

Directora:
Dra. Carolina Bonacic Castro

Línea de Investigación N° 10:
ACTUALIDAD Y DEVENIR DE LOS LENGUAJES VISUALES

Fecha de presentación:
Julio - 2026

Agradecimientos

Quiero expresar mi más sincero agradecimiento a todo el equipo del Doctorado en Diseño de la Universidad de Palermo. Especialmente a Josefina Mazzoleni y Roberto Céspedes, cuyo permanente acompañamiento humano y académico fue fundamental para llegar hasta esta etapa. Asimismo, agradezco al cuerpo académico por su compromiso, generosidad y excelencia, en particular a Roxana Ynoub, Verónica Tozzi, Daniela Fiorini, Manuel Abeledo, Karen Avenburg, Ana Cravino, Vanesa Martello y Marina Matarrese, cuyas enseñanzas, orientaciones y reflexiones contribuyeron significativamente al desarrollo de esta investigación.

A mis compañeros de doctorado, especialmente a Itzel Álvarez, Cristina Marmolejo, Laura Hurtado y Felipe Cisternas, gracias por el intercambio de ideas, la colaboración y la amistad construida durante este recorrido. Su compañía hizo de esta experiencia un proceso más enriquecedor y significativo.

De manera muy especial, agradezco a mi profesora guía, la Dra. Carolina Bonacic, por su confianza, generosidad y permanente apoyo. Más allá de sus valiosos aportes intelectuales, agradezco profundamente su calidad humana y su capacidad para inspirar, a través de su trayectoria y ejemplo, una forma comprometida de vivir la academia, promoviendo la colaboración, la innovación y una mayor participación de las mujeres en la investigación científica y tecnológica.

Finalmente, agradezco a mis amigos y a todas las personas que, de una u otra manera, contribuyeron a esta investigación con su tiempo, conocimientos y apoyo incondicional. Este trabajo también es el resultado de esos encuentros, aprendizajes y generosidades compartidas a lo largo del camino.

Índice de contenido

Capítulo I. Introducción, p. 9.

1.1. Presentación del tema. Delimitación conceptual, temporal y espacial, p. 9.

1.2. Enmarcación y justificación de la tesis dentro de las líneas temáticas de investigación de la Facultad de Diseño y Comunicación, p. 12.

1.3. Explicación de los motivos de la elección del tema de investigación, p. 14.

1.4. Relevancia de la investigación en su impacto científico o social, p. 16.

1.5. Síntesis de los capítulos, p. 17.

Capítulo II. Conceptualización, fundamentación teórica y metodología, p. 20.

2.1. Estado de la cuestión, p. 20.

2.2. Planteamiento del problema, p. 36.

2.3. Pregunta general de investigación, p. 36.

2.4. Hipótesis general, p. 37.

2.4.1. Hipótesis de trabajo, p. 37.

2.5. Objetivo general, p. 38.

2.5.1. Objetivos específicos, p. 39.

2.6. Marco teórico, p. 42.

2.6.1. De la consciencia como problema conceptual a la percepción como fenómeno de diseño, p. 50.

2.6.2. Epistemología de la consciencia, p. 54.

2.6.3. Fenomenología de la interacción, p. 61.

2.6.4. Semiótica de la interfaz, p. 68.

2.6.5. Diseño de interacción: affordance, heurística y UX conversacional, p. 91.

2.6.6. Síntesis del marco teórico para la construcción de la pauta metodológica, p. 106.

2.7. Marco metodológico, p. 109.

2.7.1. Enfoque y tipo de investigación, p. 109.

2.7.2. Diseño de investigación, p. 111.

2.7.3. Unidades de análisis y criterios de selección, p. 117.

2.7.3.1. Corpus documental: criterios de selección de textos, p. 117.

2.7.3.2. Participantes: criterio de selección por pertinencia disciplinar, p. 123.

2.7.3.3. Prueba de concepto (PoC): definición y contribución al proceso metodológico, p. 126.

2.7.4. Fases del proceso de investigación, p. 127.
2.7.5. Instrumentos y técnicas de producción de datos, p. 130.
2.7.5.1. Análisis documental: matriz de análisis, p. 130.
2.7.5.2. Entrevistas: pauta, protocolo y estrategia de análisis, p. 131.
2.7.5.3. PoC: explicación del instrumento piloto y de los criterios observados, p. 134.
2.7.6. Matriz de datos, p. 138.
Capítulo III. Análisis e interpretación de resultados, p. 144.
3.1. Levantamiento de categorías de análisis documental, p. 144.
3.2. Categorías emergentes de las entrevistas, p. 178.
3.3. Resultados de la PoC, p. 182.
3.4. Triangulación de resultados, p. 203.
3.5. Conclusiones parciales, p. 207.
Capítulo IV. Integración y sistematización de categorías para la evaluación de percepción de consciencia, p. 208.
4.1. Dimensiones de la percepción de consciencia, p. 211.
4.2. Categorías provenientes de la semiótica, p. 212.
4.3. Categorías provenientes del diseño de interacción, p. 213.
4.4. Modelo integrador, p. 214.
Capítulo V. Construcción de una pauta de evaluación metodológica, p. 217.
5.1. Dimensión: Agencia percibida, p. 217.
5.2. Dimensión: Presencia interactiva, p. 217.
5.3. Dimensión: Identidad percibida, p. 222.
5.4. Dimensión: Comprensión aparente, p. 224.
5.5. Pauta de evaluación metodológica final, p. 227.
5.6. Validación conceptual de la pauta, p. 229.
Conclusiones, p. 238.
Bibliografía, p. 242.

Índice de Tablas

Tabla 1 *Matriz de consistencia del planteamiento investigativo, p. 39*

Tabla 2 *Selección del Corpus documental, p. 118*

Tabla 3 *Perfiles de expertos por área disciplinar, p. 124*

Tabla 4 *Matriz de las fases del proceso de investigación, p. 127*

Tabla 5 *Ficha de análisis documental, p. 131*

Tabla 6 *Matriz de ámbitos temáticos de las preguntas a expertos, p. 132*

Tabla 7 *Ficha de registro de observación de la sesión - PoC, p. 135*

Tabla 8 *Instrumento de observación por dimensiones - PoC, p. 135*

Tabla 9 (parte 1) *Matriz de datos de la investigación, p. 138*

Tabla 9 (parte 2) *Estrategia de recopilación, tratamiento analítico y uso de los datos, p. 142*

Tabla 10 *Análisis comparativo de citación de fuentes en marcos conceptuales, p.145*

Tabla 11 *Diez heurísticas de Nielsen: reinterpretación para el análisis de la percepción de consciencia en chatbots, p. 152*

Tabla 12 *Ocho reglas de oro de Shneiderman: reinterpretación para el análisis de la percepción de consciencia en chatbots, p. 155*

Tabla 13 *Nielsen-Shneiderman: comparación según tipo de relación, p. 157*

Tabla 14 *Nielsen-Shneiderman: síntesis operativa de principios para la pauta de evaluación, p. 160*

Tabla 15 *Nielsen-Shneiderman: síntesis analítica de principios para la pauta de evaluación, p. 162*

Tabla 16 *Matriz de categorías Nielsen-Shneiderman orientada a la pauta de evaluación, p. 163*

Tabla 17 *Matriz conceptual del Corpus documental por pilares, p. 164*

Tabla 18 *Codificación por pilares y categorías generales, p. 167*

Tabla 19 *Matriz de Síntesis de Análisis Documental, p. 169*

Tabla 20 *Matriz primera ronda de entrevistas, p.179*

Tabla 21 *Matriz segunda ronda de entrevistas, p.180*

Tabla 22 *Matriz de criterios transversales basados en las entrevistas, p.181*

Tabla 23 *Pauta de análisis comparativo de tres casos en la PoC, p. 188*

Tabla 24 *Pauta de análisis semiótico-discursivo de ChatGPT en la PoC, p. 193*

Tabla 25 *Aplicación del instrumento de observación por dimensiones - PoC: caso ChatGPT, p.199*

Tabla 26 *Matriz de triangulación de categorías emergentes para la evaluación de percepción de consciencia, p.204*

Tabla 27 *Modelo integrador de categorías para la evaluación de percepción de consciencia, p.209*

Tabla 28 *Matriz del modelo de articulación de categorías, p.214*

Tabla 29 *Sistematización jerárquica de dimensiones, categorías, subcategorías e indicadores para la evaluación de la percepción de consciencia, p.215*

Tabla 30 (parte 1) *Pauta metodológica final: Dimensión D1. Agencia percibida, p.218*

Tabla 30 (parte 2) *Pauta metodológica final: Dimensión D2. Presencia interactiva, p.220*

Tabla 30 (parte 3) *Pauta metodológica final: Dimensión D3. Identidad percibida, p.222*

Tabla 30 (parte 4) *Pauta metodológica final: Dimensión D4. Comprensión aparente, p.225*

Tabla 30 (parte 5) *Pauta metodológica final: criterio transversal de accesibilidad, inclusión y perceptibilidad, p.227*

Tabla 31 *Validación conceptual de la pauta por triangulación de fuentes, p.229*

Tabla 32 *Pauta Final de evaluación metodológica para la percepción de consciencia en chatbots inteligentes, p.232*

Tabla 33 *Aplicación demostrativa de la pauta final de evaluación metodológica al caso ChatGPT, p.235*

Índice de Figuras

Figura 1 *Mapa conceptual del marco teórico, p. 108*

Figura 2 *Esquema proceso de investigación, p. 113*

Figura 3 *Avatar de agentes conversacionales, p. 183*

Figura 4 *Extracto de conversación con ELIZA, p. 184*

Figura 5 *Extracto de conversación con LOLA, p. 185*

Figura 6 *Extracto de conversación con ChatGPT, 2023, p. 186*

Figura 7 *Extracto de conversación con ChatGPT, 2023, consultada en 2026, p. 187*

CAPÍTULO I: INTRODUCCIÓN

1.1. Presentación del tema. Delimitación conceptual, temporal y espacial

El vínculo entre el diseño de experiencia de usuario y la interfaz de los *chatbots inteligentes* es complejo de abordar en un contexto de constante evolución tecnológica. Si bien esta investigación parte del interés por los elementos gráficos e interactivos que inciden en la percepción de consciencia en la interacción humano-máquina, el enfoque se desplaza deliberadamente desde lo empírico contingente hacia la construcción de un marco epistemológico y operativo desde el campo del diseño.

Esta decisión responde a una limitación fundamental que afecta a toda investigación que tome como objeto de estudio un sistema tecnológico en permanente transformación: una conversación de chat en un momento específico, una plataforma inmersiva, una versión digital online, un corpus conformado por capturas de pantalla queda probablemente obsoleto antes de que concluya la investigación.

En consecuencia, esta investigación busca construir un modelo integral para evaluar la consciencia percibida en *chatbots* desde el diseño de interacción, considerando la continua transformación tecnológica de estos sistemas. Para ello, se propone un conjunto de criterios semióticos, fenomenológicos e interactivos que permita analizar el fenómeno sin reducirlo al estado de desarrollo de una plataforma específica, y que pueda ser aplicado o adaptado por investigadores y diseñadores en futuros contextos de evaluación. Su adaptación y pertinencia en otros contextos constituyen posibilidades de desarrollo posterior, sustentadas en la fundamentación conceptual de la propuesta.

Con respecto al recorte espacial de la investigación y en una búsqueda inicial y exploratoria de posibles casos de estudio, se toma en cuenta que existen tantos *chatbots* como sitios web y aplicaciones móviles en Internet; considerando esta circunstancia la muestra sería miles de casos.

Por este motivo, el recorte espacial y temporal propuesto está basado en evaluar el fenómeno de la consciencia en un momento determinado, focalizando

el estudio en la experiencia de los usuarios y no en el objeto tecnológico como tal, ni en la validación con usuarios de manera directa.

Entonces, se decide la consulta a expertos del entorno más cercano de la autora, investigadores de Santiago de Chile, de universidad pública durante los años 2022 hasta inicios de 2026, de esta manera se comparte un mismo contexto sociocultural, facilitando los análisis posteriores del fenómeno.

El fenómeno visto de manera generalista, aporta a la comprensión del marco conceptual y a los factores que intervienen de manera más amplia y con menos riesgos a que la investigación quede rápidamente obsoleta. Para futuros estudios, los investigadores de este campo, contarán con un marco conceptual y metodológico robusto que pueda ser adaptado y aplicado a diversas situaciones evaluativas y contextos.

Con respecto al recorte temporal, es necesario aclarar que la investigación inició el año 2022, ya que es en ese periodo cuando hay un auge de la tecnología de la inteligencia artificial en el país. Unos meses antes, el 20 de septiembre de 2021, había sido promulgada la primera Política de Inteligencia Artificial en Chile, siendo un referente latinoamericano en materia de legislación de IA.

En paralelo, pocos meses después, el Senado chileno aprobó por votación unánime un proyecto de ley[1] que modifica la Constitución para proteger los derechos del cerebro o neuroderechos, regulando las intervenciones médicas y científicas que se pueden realizar a nivel neuronal. La senadora Carolina Goic indicó que “este proyecto es la bajada que nos permite hacer operativa la reforma constitucional de los neuroderechos. Materia en que Chile está liderando a nivel mundial dando ejemplo de cómo nos adelantamos con legislaciones que protegen a las personas" (Senado de la República de Chile, 2021, párr. 8).

Los derechos neuronales se cruzan con el tema de cómo se producen los procesos mentales y la consciencia. Con el avance de la tecnología y las diversas posturas éticas, la percepción de consciencia de un *chatbot* inteligente puede que en algunos años también tenga que ser debatido en el Senado.

[1] La norma fue propuesta el 2021 por integrantes de la Comisión de Desafíos del Futuro, Ciencia, Tecnología e Innovación del Senado de la República de Chile, específicamente por los senadores Guido Girardi, Carolina Goic, Francisco Chahuán, Juan Antonio Coloma y Alfonso De Urresti.

En Chile, producto de la pandemia, los comercios aumentaron el uso de servicios digitales. La vuelta a la normalidad comenzó gradualmente a partir del 30 de septiembre de 2021, fecha en que finalizó el Estado de Excepción Constitucional de Catástrofe, vigente desde el 18 de marzo de 2020 (Ministerio de Salud de Chile, 2021); sin embargo, los comercios continuaron mejorando sus servicios digitales. El comercio electrónico tuvo un incremento del 90% en 2022 respecto al período pre-pandémico (Cámara de Comercio de Santiago, 2024, como se citó en Quijada & Vergara, 2024), contexto en el que muchos comercios comenzaron a implementar *chatbots* como canal de atención al cliente.

El año 2022, en los comercios chilenos, comienza a ganar notoriedad la masificación de servicios digitales con *chatbots inteligentes*, en sus sitios web y aplicaciones móviles. Su primer rasgo distintivo en la interfaz era el ícono del avatar que daba inicio a la conversación. Sin embargo, en un principio no todos contaban con la incorporación de *chatbots* con inteligencia artificial, sino que sólo estaban dotados de rutinas de respuestas automáticas, previamente programadas, entregando respuestas que generalmente ocasionaban mucha frustración en el usuario, a pesar de eso, igualmente fueron parte de las conversaciones cotidianas de las personas.

En los años siguientes, *los chatbots* ya contaban con notables mejoras, tanto en el funcionamiento tecnológico, con la incorporación de la inteligencia artificial, como en el diseño de la interfaz e interacción con el usuario. En muchos casos, se han convertido en el primer contacto que tiene un cliente con un “representante” de la empresa.

Para efectos de esta investigación, se adopta el término *chatbot inteligente* como constructo operacional para designar a los sistemas conversacionales que incorporan inteligencia artificial, particularmente aquellos basados en modelos de lenguaje de gran escala y aprendizaje automático.

Esta decisión responde al desarrollo y la expansión de los *chatbots* que incorporan inteligencia artificial en contextos cotidianos, educativos, sanitarios, comerciales y corporativos. La distinción entre sistemas basados en reglas preprogramadas y sistemas con inteligencia artificial no constituye el eje central de esta investigación, aunque se considera en la selección de casos de contraste

para examinar señales de interfaz y dinámicas conversacionales en sistemas con distintas capacidades tecnológicas.

El término *chatbot inteligente* permite, por tanto, simplificar la denominación sin pérdida de precisión conceptual, al tiempo que delimita con claridad el objeto de estudio. Esta convención se aplicará de manera consistente a lo largo de todo el documento.

La delimitación conceptual, temporal y espacial presentada en este apartado revela que el fenómeno de la percepción de consciencia en *chatbots inteligentes* es, en su núcleo, un problema de lenguaje visual. Los elementos gráficos e interactivos de la interfaz, el avatar, la tipografía, los emoticones, los indicadores de escritura en curso, no son envoltorios funcionales de un sistema técnico sino configuraciones de sentido que median la experiencia del usuario y producen efectos de presencia, agencia e identidad.

En este sentido, el objeto de estudio delimitado se inscribe de manera directa en una pregunta que el campo del diseño lleva formulando desde hace décadas: ¿de qué manera los lenguajes visuales contemporáneos se hacen cargo, como huella y como poética, de los fenómenos culturales y simbólicos que atraviesan la vida de las personas? Los *chatbots inteligentes* constituyen una de las expresiones más extendidas y menos estudiadas desde el diseño de esos lenguajes en evolución.

Es por ello que la presente investigación se enmarca en la Línea de Investigación N° 10 de la Facultad de Diseño y Comunicación de la Universidad de Palermo: "Actualidad y devenir de los lenguajes visuales", tal como se desarrolla en el apartado siguiente.

1.2. Enmarcación y justificación de la tesis dentro de las líneas temáticas de investigación de la Facultad de Diseño y Comunicación.

La presente investigación doctoral, vinculada directamente con la Línea de Investigación Nº 10: “Actualidad y devenir de los lenguajes visuales” de la Universidad de Palermo, se fundamenta en:

> La línea centra su orientación en la imagen y su interpretación. Las múltiples implicancias que involucran el abordaje del polisémico concepto de frontera (en

> relación a los lenguajes visuales) son tomadas por diferentes realizadores estéticos para su dilucidación en el campo de lo sensible y lo inteligible. La preocupación que se halla en los textos y en las conferencias gira en torno al modo en que el arte, en especial la fotografía, se hizo y se hace cargo a modo de huella, relato y poética de diferentes sucesos geopolíticos entre los que se encuentran las migraciones, fenómeno que atraviesa el panorama pasado y presente mundial. (Universidad de Palermo, s.f., sección Fundamentación)

En este sentido, la tesis se inserta en las preocupaciones de la línea respecto a la interpretación de los lenguajes visuales y el análisis de la imagen como experiencia estética y cultural. Asimismo, dialoga con teorías de la semiótica visual, la comunicación, las hipermediaciones y el diseño gráfico, ampliando el estudio tradicional de la imagen hacia los entornos digitales mediados por inteligencia artificial.

En concreto, la investigación se articula con esta línea porque aborda los *chatbots inteligentes* como nuevas formas contemporáneas de lenguaje visual e interacción simbólica. En ellas, los elementos gráficos e interactivos de los *chatbots*, tales como avatares, recursos visuales, diseño de interfaz y estructuras conversacionales, no operan como meros componentes funcionales, sino como signos que producen sentido y modulan la experiencia del usuario.

El análisis sobre cómo estos elementos influyen en la percepción de consciencia durante la interacción humano-máquina se inscribe, así, en la pregunta central que recorre la Línea N° 10: de qué manera los lenguajes visuales se hacen cargo, a modo de huella y poética, de los fenómenos culturales y simbólicos que atraviesan la vida contemporánea.

De esta manera, la investigación contribuye a explorar el devenir actual de los lenguajes visuales, comprendiendo las interfaces de IA no sólo como dispositivos funcionales, sino también como construcciones visuales y discursivas capaces de producir sentido, experiencia e imaginarios sobre la consciencia artificial.

Además, esta investigación amplía la Línea N° 10 hacia un territorio emergente de los lenguajes visuales: las interfaces de los *chatbots* inteligentes. En ellas, elementos como el avatar, la tipografía, los emoticones y los indicadores de escritura no operan como componentes funcionales sino como signos que

producen sentido, modulan la experiencia del usuario y construyen una forma de presencia e identidad en la interacción humano-máquina.

La incorporación de entornos digitales mediados por inteligencia artificial no abandona la pregunta central de la línea: de qué manera los lenguajes visuales se hacen cargo de los fenómenos simbólicos y culturales de una época, sino que la lleva a uno de sus horizontes más actuales y menos explorados desde el diseño.

Si la Línea N° 10 ha interrogado la imagen fotográfica, el relato visual y la poética de los lenguajes estéticos, esta tesis propone que la interfaz conversacional es también un lenguaje visual en devenir, cuya interpretación requiere las herramientas de la semiótica visual, la fenomenología de la interacción y el diseño de interacción articuladas como marco disciplinar propio.

En síntesis, la vacancia que permite el estudio del fenómeno de percepción de consciencia en *chatbots* inteligentes aporta claramente a la Línea N° 10 de investigación en la medida que es entendida como un nuevo devenir de los lenguajes.

1.3. Explicación de los motivos de la elección del tema de investigación

La elección del tema de investigación surge del interés por comprender un fenómeno contemporáneo que se ha intensificado con el avance de la inteligencia artificial conversacional: la percepción de consciencia atribuida por los usuarios a los *chatbots inteligentes* durante la interacción humano-máquina.

Si bien este debate ha sido abordado desde distintas miradas, gran parte de los estudios existentes se han centrado en aspectos técnicos del funcionamiento de la inteligencia artificial, en la eficiencia conversacional de los sistemas o en análisis parciales de elementos específicos de la interfaz, como la redacción textual, la automatización de respuestas o la usabilidad operativa.

Esta mirada fragmentada ha generado dos problemas de fondo: por una parte, una comprensión reducida y sesgada del objeto de estudio; y por otra, la ausencia de un desplazamiento analítico desde el diseño de la interfaz hacia la interacción como fenómeno que es importante estudiar desde el diseño. De este modo, la percepción de consciencia se constituye como un problema disciplinar legítimo y necesario para el campo del diseño de interacción.

Más que comprender al *chatbot* como una herramienta tecnológica o un canal automatizado de comunicación, esta investigación propone abordarlo como un entorno semiótico e interactivo complejo, en el cual convergen signos visuales, estructuras discursivas, affordances, temporalidades de respuesta y recursos heurísticos que condicionan la experiencia fenomenológica del usuario.

El interés no reside en determinar si un *chatbot* posee consciencia real, sino en analizar cómo ciertas condiciones de diseño permiten que el usuario lo perciba como una entidad dotada de presencia, agencia o identidad.

La elección del tema responde también a la acelerada evolución tecnológica de las plataformas digitales. Los *chatbots inteligentes* cambian constantemente, incorporando capacidades que vuelven rápidamente obsoletos los análisis centrados en casos particulares o tecnologías específicas.

Frente a ello, surge la necesidad de construir un marco conceptual y metodológico transferible, capaz de mantenerse vigente más allá de la contingencia tecnológica.

Desde la noción de vacancia cognitiva, esta investigación resulta pertinente, ya que existe un vacío disciplinar respecto a la construcción de criterios sistemáticos que permitan analizar y evaluar la percepción de consciencia en *chatbots inteligentes* desde el diseño.

Actualmente, muchas de las características presentes en los *chatbots* contemporáneos, como la coherencia conversacional, la personalización dinámica, la simulación de empatía, la construcción visual de identidad, la interacción multimodal y la persistencia discursiva, exceden las categorías tradicionales utilizadas para estudiar la interacción humano-computador y no pueden ser capturadas con los instrumentos analíticos disponibles en la investigación en diseño.

Este estado de vacancia permite profundizar no sólo en la comprensión teórica del fenómeno, sino también en el desarrollo de herramientas metodológicas capaces de articular miradas provenientes de la semiótica visual, la fenomenología de la interacción y la heurística del diseño de interfaz. A su vez, esta articulación posibilita ampliar las perspectivas tradicionales sobre el diseño de interfaces digitales, desplazando el análisis desde la funcionalidad técnica

hacia la construcción simbólica de la experiencia interactiva y hacia la sistematización de criterios evaluativos que trasciendan lo situado y lo contingente.

1.4. Relevancia de la investigación en su impacto científico o social

El campo de estudio vinculado a la percepción de consciencia en *chatbots* inteligentes posee una creciente relevancia científica y social debido al acelerado avance de las tecnologías conversacionales y a su progresiva incorporación en múltiples ámbitos de la vida cotidiana, educativa, laboral y comunicacional.

En este contexto, la presente investigación busca contribuir a la comprensión de un fenómeno que, si bien ha sido abordado desde disciplinas como la filosofía, la ingeniería y las ciencias cognitivas, aún presenta importantes vacíos respecto al papel que desempeña el diseño de interacción en la construcción de dichas percepciones.

El principal aporte científico de esta tesis radica en ampliar las perspectivas tradicionales sobre la interacción humano-computador, desplazando el análisis desde enfoques predominantemente técnicos hacia una comprensión integral de la experiencia interactiva. En este sentido, la investigación propone comprender los *chatbots* inteligentes no sólo como herramientas funcionales, sino como entornos semióticos complejos en los cuales intervienen elementos visuales, discursivos, temporales e interactivos capaces de influir en la atribución de presencia, agencia o identidad por parte del usuario.

A nivel profesional y aplicado, los potenciales beneficiarios de esta investigación corresponden a diversos actores vinculados al diseño y desarrollo de experiencias digitales, entre ellos diseñadores UX/UI, publicistas, programadores, periodistas, investigadores, editores de contenidos, desarrolladores de inteligencia artificial e ingenieros, tanto con formación formal como provenientes de espacios interdisciplinarios. La investigación puede contribuir a fortalecer la comprensión de los procesos teóricos, metodológicos y simbólicos involucrados en la interacción humano-*chatbot*, permitiendo desarrollar criterios más reflexivos y críticos para el diseño de *chatbots* inteligentes contemporáneos.

Finalmente, la novedad de esta investigación es su aporte al campo del diseño de interacción mediante la construcción de una pauta metodológica de evaluación adaptable a futuros contextos tecnológicos. Esta propuesta busca constituirse como una herramienta transferible capaz de orientar el análisis crítico de *chatbots* inteligentes más allá de tecnologías específicas o contextos contingentes, contribuyendo tanto al ámbito académico como al ejercicio profesional en áreas vinculadas al diseño digital, la comunicación y la experiencia de usuario.

1.5. Síntesis de los capítulos

La investigación se inicia con el Capítulo I, donde se presenta y delimita el objeto de estudio: la percepción de consciencia atribuida a los *chatbots* inteligentes desde la perspectiva del diseño de interacción.

En este apartado se establece el recorte conceptual de la investigación, centrado en los elementos gráficos e interactivos de los *chatbots* inteligentes y en los efectos de agencia, presencia e identidad que estos producen en los usuarios.

Asimismo, se define el recorte temporal comprendido entre los años 2022 hasta mayo de 2026, período marcado por la expansión de la inteligencia artificial conversacional, y el recorte espacial correspondiente al contexto académico y sociocultural de Santiago de Chile, donde se desarrollan las entrevistas a expertos.

El capítulo también justifica la pertinencia de la investigación dentro de la Línea N.º 10 “Actualidad y devenir de los lenguajes visuales” de la Universidad de Palermo, expone los motivos que originan el estudio y fundamenta su relevancia científica, disciplinar y social.

El Capítulo II desarrolla la conceptualización, fundamentación teórica y metodología de la investigación. Inicia con el Estado de la Cuestión examinando las principales investigaciones sobre inteligencia artificial conversacional, antropomorfización de agentes, diseño de identidad visual, UX conversacional, percepción de agencia y aplicaciones de la semiótica a interfaces digitales, identificando las limitaciones de los enfoques actuales y la vacancia disciplinar que justifica esta tesis.

Posteriormente, se presentan el problema de investigación, las preguntas, hipótesis y objetivos que orientan el estudio. A continuación, el Marco Teórico articula cuatro pilares de análisis: epistemológico, fenomenológico, semiótico y de diseño de interacción, que permiten comprender la percepción de consciencia como un fenómeno emergente de la interacción entre usuarios y sistemas conversacionales.

A continuación, el Marco Metodológico describe el enfoque cualitativo adoptado, las unidades de análisis, los criterios de selección documental, la selección de expertos, la prueba de concepto y las técnicas de producción y análisis de datos utilizadas en la investigación.

El Capítulo III está dedicado al análisis e interpretación de los resultados obtenidos mediante el análisis documental, las entrevistas a expertos y la prueba de concepto. En esta etapa se identifican, contrastan y sistematizan las categorías emergentes provenientes de las distintas fuentes de información, examinando las convergencias y divergencias de los hallazgos.

El objetivo central de este capítulo es construir una base analítica sólida que permita reconocer los factores semióticos, fenomenológicos e interactivos que intervienen en la percepción de consciencia atribuida a los *chatbots* inteligentes.

El Capítulo IV sistematiza las categorías obtenidas durante el proceso analítico y las organiza en una estructura conceptual coherente para su posterior operacionalización.

En este apartado se integran las categorías provenientes de la semiótica de la interfaz y del diseño de interacción, estableciendo relaciones entre ellas y definiendo las dimensiones que servirán de fundamento para la construcción de la pauta metodológica propuesta por la investigación.

El Capítulo V presenta la principal contribución original de la tesis: la construcción de una pauta de evaluación metodológica orientada al análisis de la percepción de consciencia en *chatbots* inteligentes. Para ello, se examinan las categorías identificadas desde los pilares epistemológico, fenomenológico, semiótico y del diseño de interacción, transformándose en dimensiones, criterios e indicadores de evaluación.

El resultado es una propuesta metodológica concebida para ser transferible, adaptable a distintos contextos tecnológicos y aplicable a futuras generaciones de sistemas conversacionales inteligentes.

Finalmente, las conclusiones integran los principales hallazgos de la investigación, evalúan el cumplimiento de los objetivos e hipótesis planteados y discuten los aportes teóricos y metodológicos realizados al campo del diseño de interacción. Asimismo, se reflexiona sobre las posibilidades de aplicación futura de la pauta propuesta y sobre nuevas líneas de investigación vinculadas a la percepción de consciencia, la semiótica de la interfaz y las experiencias de interacción humano-tecnología.

Capítulo II: Conceptualización, fundamentación teórica y metodología

2.1. Estado de la cuestión

El estudio de la percepción de consciencia en *chatbots* ha sido cuestión de discusión por parte de los autores aunque más bien de manera indirecta, con foco en la agencia de las máquinas, es decir, si son capaces de tomar decisiones o si las emociones y la creatividad son una forma indirecta de inferir que detrás de una interfaz conversacional hay una mente consciente. ¿Las máquinas pueden ser conscientes? ¿cómo se puede analizar y evaluar la consciencia en máquinas? y ¿cuál es la mirada disciplinar más próxima para estudiar este fenómeno? Este ha sido el debate que ha enmarcado desde sus inicios a esta forma de comunicarse por intermedio de *chatbots* inteligentes que se ha vuelto cada vez más habitual en la vida de las personas.

El debate de los autores sobre la consciencia artificial durante la última etapa del siglo pasado ha estado alejado del diseño de interacción, ya que se ha enfocado, más bien, en las características técnicas, en los elementos gráficos antropomorfizados, en los modos de escritura y en otros aspectos particulares más que en el fenómeno mismo de la percepción de consciencia artificial y en los mecanismos para identificarla y evaluarla.

El debate más próximo al diseño de interacción y percepción de consciencia de sistemas artificiales se sitúa durante la década del noventa, específicamente en el área de la heurística y el diseño de interfaces, en tanto fue examinada su relación con la respuesta social automática de los usuarios ante sistemas computacionales y con los primeros intentos de sistematizar criterios evaluativos de usabilidad que, reinterpretados desde el diseño de interacción, constituyen antecedentes directos de las categorías analíticas de esta investigación.

A partir de lo anteriormente expuesto, el estado de la cuestión de esta tesis doctoral se organiza en seis líneas temáticas, estructuradas de lo general a lo específico. Las primeras líneas mapean el campo amplio de la IA conversacional y la interacción humano-máquina y las líneas finales se aproximan al objeto de estudio desde la UX, el diseño del agente y la percepción de consciencia (o lo

más cercano a ello); posteriormente, para cerrar este apartado, se identifican las limitaciones de los enfoques actuales y la vacancia disciplinar que justifica el aporte original de esta investigación.

Las fuentes del estado de la cuestión que se presentan a continuación, fueron seleccionados principalmente de artículos indexados en Scopus o Web of Science (WoS). Considerando la dificultad para encontrar publicaciones específicas del tema, se amplió el campo de búsqueda a artículos publicados entre 2015 y 2025. Además, se incluyen algunas monografías de referencia cuando son fundacionales (Reeves y Nass, 1996; Turkle, 2011). Se excluyeron algunos autores que forman parte del marco teórico (Peirce, Eco, etc.), mencionados en el estado de la cuestión sólo como antecedentes de la vacancia interdisciplinar.

La primera línea establece el mapa general del campo, reúne investigaciones contemporáneas sobre IA conversacional y *chatbots*, desde distintas disciplinas y con determinados énfasis, aportando un contexto amplio que más adelante justifica la relevancia del problema de investigación.

Los temas que se abordan están relacionados con la evolución tecnológica de los *chatbots*: de los sistemas basados en reglas (ELIZA, ALICE) a los modelos de lenguaje de gran escala (LLM: GPT, Gemini, Claude), con la expansión del uso de *chatbots* en contextos como los servicios, salud, educación y acompañamiento social. También se presenta el panorama interdisciplinar de la investigación: informática, psicología, comunicación, marketing y ciencias del comportamiento; y tendencias de investigación: revisiones sistemáticas de la literatura sobre interacción humano-computadora (HCI).

Se inicia esta primera línea con un mapa del campo imprescindible construido por Rapp, Curti y Boldi (2021) que estudiaron el lado humano de la interacción entre humanos y *chatbots*. Realizaron una revisión exhaustiva y sistemática de la literatura por diez años (2009-2019) sobre *chatbots* basados en texto. Los autores identificaron que la humanidad percibida en el *chatbot* no es un atributo fijo del sistema sino un fenómeno emergente que depende simultáneamente del contexto de interacción, los objetivos del usuario, y las

señales (*cues*) que el *chatbot* exhibe como signo, pero sin vincularlo con la semiótica y sin identificar que son exactamente los representámenes.

En cuanto a la evaluación, Rapp, Curti y Boldi (2021) señalaron como limitación crítica del campo que la mayoría de los estudios son experimentales y cuantitativos, y que faltan investigaciones orientadas al diseño con fundamento teórico que articulen percepción del usuario, señales de la interfaz y principios de diseño.

Por otra parte, se presenta el panorama técnico y disciplinar del campo en la investigación de Adamopoulou y Moussiades (2020) que ofrecieron una visión general de la expansión del uso de *chatbots* en contextos de servicio, salud, educación y acompañamiento social. Señalaron que los *chatbots* tienen ventajas claras tanto para usuarios como para desarrolladores (disponibilidad inmediata, plataforma independiente sin necesidad de instalación, escasa carga de diseño para la interfaz), pero que esta facilidad de desarrollo no ha sido acompañada por comprensión del impacto en el usuario.

Adamopoulou y Moussiades (2020) constataron que la confianza del usuario en el *chatbot* depende de señales de diseño relacionadas con la semejanza humana del sistema, entre ellas cómo se autopresenta, qué tan profesional es su apariencia y en qué medida responde como un humano. Esta tripartición visual, identitaria y conversacional constituye, sin que los autores lo formulen en términos semióticos (no como un sistema de señales que activan en el usuario esquemas de atribución de agencia), caracterizaron al *chatbot* como una entidad que puede funcionar simultáneamente como herramienta, juego y compañero, descripción que corresponde a tres regímenes distintos de percepción de consciencia y cuya modulación depende precisamente del diseño de la interfaz.

Følstad y Brandtzaeg (2017) abordaron la interacción persona-ordenador desde la mirada interdisciplinaria: informática, psicología, comunicación, marketing, ciencias del comportamiento. Propusieron un giro conversacional en HCI al estudiar *chatbots* y plantearon la interacción como un nuevo mundo, el usuario ya no navega un sistema sino que conversa con un agente. Esa reformulación tiene consecuencias para la percepción en el modo conversacional,

el usuario activa esquemas cognitivos y sociales propios de la interacción interpersonal, que incluyen la atribución de intencionalidad, empatía y agencia al interlocutor.

Dentro del paradigma gráfico, Følstad y Brandtzaeg (2017) concluyeron que el *chatbot* no es simplemente una nueva interfaz sino una ruptura de paradigma, la interacción ya no ocurre a través de elementos visuales que el usuario manipula sino a través de cadenas de texto en lenguaje natural.

Piccolo et al. (2019) avanzaron hacia un marco para evaluar los sistemas de *chatbot*, presentando tendencias de investigación a partir del análisis de la literatura sobre interacción humano-computador (HCI). La revisión sistematizó 131 publicaciones de la *ACM Guide to Computing Literature* entre 1975 y 2018, donde identificaron que los atributos cualitativos percibidos por el usuario, como la simpatía (*likeable*), la cortesía (*politeness*), la personalidad (*personable*), la empatía (*empathy*) y la confiabilidad (*trustworthy*), comenzaban a aparecer en la literatura como dimensiones evaluables, pero que las herramientas existentes en 2018 eran de naturaleza técnica o funcional y no contaban con marcos adecuados para evaluarlos.

En síntesis, en la primera línea se observa que la literatura del campo es predominantemente técnica (NLP, rendimiento del sistema) o conductual (satisfacción, intención de uso). La dimensión del diseño gráfico y semiótico de la interfaz como productor de percepción de consciencia no aparece como línea de investigación consolidada.

La segunda línea está focalizada en antropomorfización de agentes conversacionales e IA, siendo la línea más directamente vinculada con la hipótesis central de la tesis: el fenómeno de atribuir características humanas a sistemas artificiales. Al mismo tiempo, constituye el antecedente empírico que de mejor manera se acerca al problema de investigación.

Los temas que se abordan están relacionados con las definiciones y mediciones del antropomorfismo en agentes de IA: atribución de experiencia, agencia, consciencia y semejanza humana; el paradigma CASA (*Computers Are Social Actors*) que estudia de manera empírica la respuesta social automática de los computadores; el antropomorfismo y confianza: efectos del diseño

antropomórfico en la percepción de inteligencia, empatía artificial y satisfacción del usuario; y el antropomorfismo como espada de doble filo: humanizar el sistema genera engagement o por el contrario, genera escepticismo o rechazo. Además se aborda la atribución de consciencia e intencionalidad en agentes de IA: *mind perception, theory of mind* aplicada a interfaces.

Esta segunda línea inicia con los aportes de Reeves y Nass (1996), constituyéndose como un antecedente empírico fundacional sobre los medios (computadores, televisión y nuevos medios), abordaron el paradigma CASA, concluyendo que no es que los usuarios crean que los computadores son humanos sino que responden *como si* lo fueran, aunque sepan que no lo son. Un hallazgo especialmente relevante es que las señales mínimas visuales son suficientes para activar respuestas sociales en el usuario.

En el prólogo del libro de Fogg, Zimbardo señaló que la investigación de Fogg exploró los paralelos entre la influencia persuasiva de los computadores y los agentes humanos, ampliando el trabajo de Reeves y Nass sobre las formas en que las personas responden a los computadores y otros medios como si fueran personas reales (Zimbardo, citado en Fogg, 2003, p. xi, trad. propia).

En complemento, el estudio de Gambino et al. (2020) ampliaron el paradigma CASA, realizando una actualización crítica que responde a los cambios ocurridos en las personas, las tecnologías y las formas de interacción desde los años noventa. Los autores argumentaron que los humanos han desarrollado guiones específicos de interacción con los medios, distintos de los guiones humano-humano que el CASA original postulaba. Cuando un usuario interactúa con un *chatbot* de manera no deliberada, no aplica mecánicamente las mismas normas sociales que usaría con otra persona, sino guiones propios de la interacción humano-medio que han sido aprendidos culturalmente.

Junto a lo anterior, Guingrich y Graziano (2025) examinaron la relación entre el uso de *chatbots* de compañía, el antropomorfismo y las relaciones sociales humanas. En un estudio longitudinal con 183 participantes asignados aleatoriamente a interactuar con un chatbot o realizar juegos de palabras durante 21 días, no encontraron cambios estadísticamente significativos en la salud social y las relaciones humanas atribuibles al uso del chatbot, en comparación con el

grupo de control. Sin embargo, una mayor atribución de características humanas al agente se asoció con un mayor impacto percibido de la interacción en las relaciones con familiares y amigos. Estos resultados son relevantes para el diseño de interfaces porque invitan a examinar las condiciones que favorecen el antropomorfismo y su relación con la experiencia social del usuario, sin establecer por sí solos efectos causales de las señales gráficas e interactivas ni consecuencias sostenidas a largo plazo.

Araujo (2018) analizó la influencia de las señales de diseño antropomórficas y el encuadre de la agencia comunicativa en cuanto a las expectativas generadas por los *chatbots*: diseño antropomórfico, presencia social y conexión emocional. El estudio dostró que señales conversacionales de tipo humano, como el estilo del lenguaje y el nombre del agente, aumentaron tanto el antropomorfismo consciente (*mindful*) como el inconsciente (*mindless*), y que la presencia social así generada actuó como mediador de la conexión emocional del usuario con el sistema.

Por otra parte, Epley et al. (2007) propusieron un marco teórico de referencia sobre los mecanismos cognitivos del antropomorfismo, distinguiendo una teoría de tres factores: el conocimiento antropocéntrico accesible y aplicable (*elicited agent knowledge*), la motivación de eficacia para interactuar efectivamente con agentes del entorno (*effectance motivation*), y el deseo de contacto social y afiliación (*sociality motivation*). Esta teoría, conocida como modelo SEEK, predice que el usuario es más propenso a antropomorfizar cuando el conocimiento sobre el comportamiento humano es más accesible, cuando está motivado a predecir y controlar el comportamiento del agente, y cuando experimenta una necesidad de conexión social.

Con una visión un poco más allá, Turkle (2011) aportó una perspectiva sociológica sobre los vínculos afectivos que se construyen con sistemas digitales. A partir de la investigación etnográfica con usuarios de robots y sistemas digitales, Turkle describió el “momento robótico": el instante en que el usuario suspende la incredulidad y trata al sistema como una entidad con vida interior. Este fenómeno, que la autora observó especialmente en usuarios solitarios o en situaciones de baja conexión social, es directamente pertinente para la dimensión de la *sociality*

motivation del modelo de Epley et al. (2007) y para el concepto de *alteridad digital* que se desarrollará más adelante en el marco fenomenológico de esta tesis.

En síntesis, en la segunda línea se observa que los estudios de antropomorfismo se centran en las variables psicológicas y conductuales de la respuesta del usuario, y en los efectos del lenguaje natural. El rol específico de los elementos gráficos de la interfaz (avatar, color, tipografía, iconografía) como productores de antropomorfismo ha recibido una atención difusa y no de manera sistemática.

La tercera línea sobre UX conversacional y diseño de agentes, aborda temas relacionados con principios de diseño de la interacción conversacional: *turn-taking*, gestión del error, personalidad del agente, naturalidad discursiva. También aborda metodologías de evaluación de *chatbots* desde la UX: escalas de usabilidad conversacional, evaluación heurística adaptada. Diseño de la personalidad del agente: coherencia identitaria, tono, registro lingüístico y su efecto en la percepción del usuario.Y también, UX conversacional y presencia social: cómo el diseño de la conversación produce efectos de presencia y agencia. *Chatbot Usability Scale* y otros instrumentos de evaluación específicos.

En este punto es necesario aclarar que esta línea articula el campo anterior con el siguiente, introduciendo los criterios y metodologías de diseño de la interacción conversacional, junto con posicionar el diseño del agente como variable independiente de la experiencia del usuario. Es la bisagra entre la investigación sobre antropomorfismo (línea dos) y el diseño específico de la identidad visual del *chatbot* (línea cuatro).

La línea inicia con los aportes de Borsci et al. (2022), quienes desarrollaron y realizan una prueba piloto de dos instrumentos para evaluar la interacción con chatbots: una lista de cotejo diagnóstica (BOT-Check) y un cuestionario de 15 ítems denominado *Bot Usability Scale* (BUS-15). Este último evalúa la satisfacción con la experiencia de interacción mediante cinco factores: accesibilidad percibida, calidad percibida de las funciones del chatbot, calidad de la conversación y de la información proporcionada, privacidad y seguridad percibidas, y tiempo de respuesta.

Por su parte, Gnewuch et al. (2017) abordaron los principios de diseño conversacional en el diseño de agentes conversacionales cooperativos y sociales para el servicio al cliente. Los autores identificaron que el diseño del agente como entidad social, y no solo como sistema de respuesta automática, incide directamente en la satisfacción y la percepción de utilidad del usuario. Además, propusieron principios de diseño destinados a producir cooperación social.

En complemento con lo anterior, Luger y Sellen (2016) estudiaron la brecha entre las expectativas del usuario y la experiencia de los agentes conversacionales. A través de entrevistas con usuarios de asistentes de voz como *Siri* y *Cortana*, los autores documentaron que los usuarios construyen modelos mentales del sistema basados en metáforas humanas, lo que genera expectativas que el sistema frecuentemente no puede satisfacer. Cuando esas expectativas se incumplen, el usuario experimenta frustración, y lo atribuye no a sus propias expectativas sino a una promesa rota del sistema.

En cuanto al diseño y a la evaluación de *chatbots* basados en texto, Mafra et al. (2024) desarrollaron una propuesta de tecnologías de usabilidad-UX, constituyéndose como un marco específico de evaluación UX. Los autores propusieron una serie de métodos de evaluación adaptados a las particularidades de los *chatbots* de texto, incluyendo pruebas de tarea, cuestionarios de satisfacción y análisis de transcripciones conversacionales; sin embargo, la propuesta se mantuvo centrada en el rendimiento funcional y en la satisfacción, sin abordar la dimensión semiótica ni la percepción de consciencia.

En síntesis, en la tercera línea se encuentran pautas de evaluación e investigación empírica con chatbots que se acercan a las dimensiones sociales y de percepción de usuario, sigue siendo predominante el enfoque en el rendimiento y la funcionalidad del sistema. La dimensión semiótica del diseño de la interfaz como productora de efectos de presencia no está integrada en los marcos de evaluación existentes. Este límite ratifica la vacancia que esta tesis busca cubrir desde el diseño de interacción.

La cuarta línea, sobre el diseño de la identidad visual y la personalidad de los *chatbots*, investiga específicamente el efecto del diseño visual (avatar, nombre, tipografía, color, iconografía) y su impacto en la percepción del usuario.

Los temas específicos que se abordan en esta línea están relacionados con el efecto del avatar en la percepción de empatía, confianza y consciencia del *chatbot*: avatares humanos, caricaturizados, abstractos y sin avatar. La identidad singular vs. identidad genérica del agente, la atribución de género y raza, elementos de connotación de inteligencia y acción.

Una investigación reciente de Chen et al. (2024) examinó los efectos de la apariencia antropomórfica y del estilo conversacional de los *chatbots* mediante seguimiento ocular (*eye tracking*) y cuestionarios, con datos válidos de 76 participantes que observaron prototipos con diálogos en pantalla. Encontraron que los avatares humanos y animados atrajeron mayor atención visual en el área de su representación que una burbuja de diálogo, mientras que la apariencia antropomórfica y el estilo conversacional humanizado favorecieron la presencia social percibida, la confianza y la satisfacción. Estos resultados aportan evidencia empírica sobre la influencia de las señales visuales y conversacionales en la valoración del agente, sin medir directamente la atribución de agencia o consciencia.

Cao et al. (2025) exploraron cómo cuatro tipos de avatar (humano, caricaturizado, zoomórfico y funcional) influyen en las primeras impresiones de los usuarios sobre agentes de atención al cliente. El estudio examinó la relación entre la apariencia visual, la identificación del agente como humano o IA y las valoraciones de los participantes.

Mediante seguimiento ocular y evaluaciones subjetivas, con datos válidos de 56 participantes, encontraron que los avatares humanos y caricaturizados favorecen la calidez, la competencia, la presencia social y la confianza percibidas. Las apariencias más humanas también aumentaron la probabilidad de que los participantes identifiquen al agente como humano. El estudio aporta evidencia sobre la relevancia de la apariencia visual, aunque se basó en la observación de avatares, sin interacción conversacional activa.

Aunque sus resultados no respaldaron la hipótesis del valle inquietante (*uncanny valley*), el estudio la aborda desde la apariencia visual del avatar. En la siguiente línea, este concepto será retomado desde la perspectiva de la percepción cognitiva.

En complemento con lo anterior, Song y Shin (2024) también investigaron sobre los efectos del valle inquietante (*uncanny valley*) y familiaridad del avatar, en cuanto a la percepción de la confianza en los *chatbots*, en la intención de compra y en la intención de adopción. Los autores encontraron que el efecto *uncanny valley* en *interfaces conversacionales* está moderado por la familiaridad previa del usuario con el avatar: un avatar que el usuario ya conoce puede producir menor incomodidad aunque tenga un alto grado de semejanza humana.

Para Go y Sundar (2019), las señales visuales, de identidad y conversacionales intervienen de manera diferenciada y combinada en la experiencia con agentes conversacionales. En un experimento con 141 participantes, encontraron que una mayor interactividad de los mensajes favoreció la presencia social percibida y pudo compensar una apariencia visual poco antropomórfica, mientras que presentar al agente como humano elevó las expectativas sobre su desempeño conversacional.

En este mismo sentido, Gnewuch et al. (2018) investigaron sobre las señales visuales de acción: el efecto del indicador de escritura ("...") en la percepción de presencia social. Los estados del sistema y su efecto en la percepción de deliberación y presencia demostraron que una demora dinámica en la respuesta del *chatbot*, simulada por el indicador de escritura en curso, produjo mayor percepción de presencia social que una respuesta inmediata, siendo estrecha la percepción de naturalidad y variable culturalmente.

Por otra parte, Umair et al. (2026) investigaron cómo la definición y la declaración explícita de la identidad de los *chatbots* influyeron en la autenticidad percibida, la comodidad en la interacción, la evaluación del servicio y el apego a la marca. Los autores abordaron la relación entre calidez y transparencia mediante tres formas de presentación del agente: humana, artificial e híbrida, esta última combinando un trato humanizado con la declaración de su naturaleza artificial. En un experimento con 1.200 participantes que revisaron diálogos simulados en contextos de comercio electrónico y hotelería, los autores encontraron que la identidad híbrida obtuvo las valoraciones más favorables en las dimensiones evaluadas. Para esta tesis, este antecedente permite distinguir la identidad singular de la identidad genérica del agente y plantea la pertinencia de examinar

cómo la asignación de un nombre al chatbot, la consistencia de su personalidad y las señales de género y raza pueden intervenir en el vínculo usuario-sistema.

En síntesis, en la cuarta línea se observa que los estudios sobre identidad visual del *chatbot* son predominantemente experimentales y cuantitativos. No se encontró un marco semiótico sistemático que permita analizar y evaluar los elementos gráficos de la interfaz como un sistema de signos productor de consciencia percibida.

La quinta línea aborda temas sobre la percepción de agencia, confianza e intencionalidad en sistemas de IA, en los mecanismos cognitivos y perceptuales por los cuales los usuarios atribuyen estas características a los sistemas de IA. Es el sustento empírico más directo de la hipótesis central de la tesis.

Esta línea inicia con los aportes de Gray et al. (2007) que desarrollaron un marco de referencia conceptual, distinguiendo dos dimensiones de la percepción de la mente: experiencia y agencia, donde la percepción de consciencia en *chatbots* puede involucrar una o ambas dimensiones, por ejemplo, provocar la capacidad de sentir, desear, emocionarse (experiencia) y la capacidad de planear, actuar, controlar (agencia).

Por su parte, Waytz et al. (2010) investigaron sobre la estabilidad y la importancia de las diferencias individuales en la tendencia a antropomorfizar, al correlacionar variables como la necesidad de conexión social y la tendencia a percibir agencia en el entorno, observaron que el diseño de la interfaz opera sobre tendencias variables pero percibidas como reales.

En este punto se retoma el efecto de *uncanny valley* visto en la línea anterior desde la perspectiva del diseño del avatar, en cambio, ahora es analizado desde la perspectiva de la percepción cognitiva; Cihodaru-Stefanache y Podina (2025) realizaron una revisión sistemática sobre *uncanny valley* en agentes conversacionales, evidenciando que si bien el fenómeno es real varía según el contexto de uso, el tipo de tarea y la cultura del usuario. El hecho de que el efecto *uncanny valley* esté presente en dos líneas desde distintas perspectivas, demuestra la variabilidad y complejidad del fenómeno.

Desde otro punto de vista, Westerman et al. (2019) investigaron sobre la percepción de humanidad en interacciones textuales que simulan una

conversación con un *chatbot*, examinando qué señales llevan a los usuarios a percibir al interlocutor como más o menos humano, sin avatar ni voz. En un experimento con 83 participantes, compararon respuestas con y sin errores tipográficos y palabras en mayúsculas. Encontraron que los errores tipográficos redujeron la humanidad percibida y la sensación de cercanía con el interlocutor, mientras que las mayúsculas no produjeron efectos estadísticamente significativos. Para esta tesis, estos resultados justifican examinar otras señales discursivas, como la fluidez lingüística y la capacidad de reparar malentendidos, cuya relación con la humanidad percibida requiere evidencia específica.

A su vez, Cheng et al. (2022) se interesaron por comprender el impacto del antropomorfismo en la respuesta del consumidor desde la perspectiva de la confianza y las normas relacionales, encuentraron que el antropomorfismo activa preferentemente normas centradas en el cuidado mutuo, lo que aumentaron la confianza en el sistema, pero también elevó las expectativas de reciprocidad emocional.

En síntesis, en la quinta línea se observa que los estudios de percepción de agencia son predominantemente psicológicos y no abordan el rol del diseño gráfico como variable independiente. La relación entre los signos visuales de la interfaz y la percepción de agencia no ha sido sistematizada desde un enfoque de diseño de interacción.

La última línea, la sexta, trata sobre la semiótica aplicada a *chatbots inteligentes*, presentando estudios que han aplicado marcos semióticos al análisis de interfaces, avatares y agentes conversacionales. Es la línea más cercana al aporte metodológico original de esta investigación, presenta un mapeo de los antecedentes directos de el enfoque metodológico de esta tesis puede considerar.

La búsqueda sobre semiótica visual en *chatbots* se realizó en Scopus en el rango temporal de 2018-2025 por "*semiotic analysis chatbot interface", "visual semiotics conversational agent";* esta línea de investigación es la que presenta un mayor vacío bibliográfico específico, algunas temáticas encontradas son: semiótica de la interfaz (análisis del sistema de signos de las interfaces digitales), semiótica del avatar (como signo complejo productor de significados), retórica visual (aplicada al diseño de interfaces conversacionales), análisis semiótico de

elementos gráficos (color, tipografía e iconografía en interfaces de *chatbots*), marcos integrados de la relación entre semiótica en HCI (análisis de la experiencia de usuario).

Desde la fotografía, Sonesson (1994) investigó sobre la semiótica del signo visual, aporta como antecedente al análisis del avatar la noción de índice visual en la fotografía, y sobre la relación entre el signo visual y su referente, ofrece herramientas conceptuales para pensar el avatar del *chatbot* como un signo icónico-indicial que produce efectos de presencia y referencialidad. Para esta investigación doctoral, el aporte del autor no es directo sino genealógico: establece una línea de análisis de la imagen visual como signo complejo que esta tesis hereda y aplica al contexto de los *chatbots* inteligentes, donde el avatar no representa a una persona real sino que construye una identidad virtual.

Desde los medios digitales, Nöth (1997) investigó sobre la semiótica, aportando un panorama conceptual, su trabajo sistematizó los principales enfoques semióticos aplicados al análisis de medios, incluyendo la semiótica del texto digital, de la imagen y de la interfaz. Esta investigación es importante ya que el autor proporciona un marco de referencia que muestra que la semiótica tiene una tradición de aplicación a los medios digitales anterior a esta tesis, pero que esa tradición no se ha extendido de manera específica al análisis de *chatbots inteligentes* ni a la percepción de consciencia en *chatbots*. Este vacío de aplicación específica es uno de los argumentos de la vacancia disciplinar que esta investigación busca cubrir.

Desde la ingeniería, de Souza (2005) aporta un marco semiótico específico para HCI; constituyéndose como una referencia, su propuesta de ingeniería semiótica (*semiotic engineering*) concibe el proceso de diseño de interfaces como un acto de comunicación entre diseñadores y usuarios: el diseñador codifica en la interfaz un mensaje sobre cómo debe ser usada, y el usuario lo decodifica durante la interacción. Esta perspectiva tiene una conexión directa con el posicionamiento semiótico-constructivista de esta tesis, pero también una diferencia sustantiva: de Souza centra el análisis en la comunicabilidad del diseño (si el usuario comprende lo que el diseñador quiso decir), mientras que esta investigación centra el análisis en la capacidad de la interfaz para producir percepción de consciencia (si el

sistema de signos activa en el usuario la atribución de agencia). Son problemas distintos que utilizan herramientas semióticas similares pero con objetivos diferentes.

En complemento con lo anterior, Prates et al. (2000) investigaron sobre un método para la evaluación comunicacional de interfaces de usuario desde la semiótica mediante el análisis de transcripciones de interacción, sin embargo, evaluando sólo la función.

En síntesis, en la sexta línea se observa que la semiótica aplicada a HCI existe como campo: *semiotic engineering* en Prates et al. (2000), pero su aplicación al análisis y evaluación del *chatbot* como productor de consciencia percibida es prácticamente inexistente en la literatura indexada. Los estudios que combinan semiótica y *chatbots* inteligentes son escasos, fragmentarios y no articulan de manera sistemática las tres tradiciones semióticas centrales de esta tesis (Peirce, Eco, Groupe μ) con el fenómeno de la percepción de agencia.

A modo de síntesis general del Estado de la Cuestión se identificaron varias limitaciones que justifican esta investigación doctoral:

El predominio del enfoque cuantitativo-experimental en la mayoría de los estudios sobre percepción de consciencia y antropomorfismo en *chatbots*. Se utilizan diseños experimentales que miden variables conductuales (intención de uso, satisfacción, confianza) sin analizar el sistema de signos de la interfaz como variable independiente. Los signos son tratados como estímulos, no como sistemas semióticos estructurados.

Los enfoques existentes son insuficientes para sistematizar desde la literatura una metodología de análisis y evaluación de los elementos gráficos de la interfaz del *chatbot* como productores de consciencia percibida.

La fragmentación disciplinar de los estudios de HCI, psicología cognitiva, comunicación y marketing. Se aborda el fenómeno desde sus propias categorías sin integración teórica. No se encontró en el corpus consultado un marco transdisciplinar que articule semiótica, fenomenología y diseño de interacción para el análisis del *chatbot*.

La subordinación del diseño a la ingeniería. Se detectó que la investigación sobre *chatbots* es liderada por áreas de informática, NLP y psicología. El diseño

de interacción participa como aplicación de criterios definidos por otros campos, no como disciplina con marcos propios de análisis y evaluación.

En este punto resulta pertinente retomar la contribución de Følstad et al. (2021), quienes, en el artículo Future directions for *chatbot* research: An interdisciplinary research agenda, identificaron una serie de desafíos y orientaciones futuras para la investigación sobre *chatbots* desde una perspectiva interdisciplinaria.

La ausencia de una teoría del signo aplicada al *chatbot*. Los estudios que abordan el diseño visual del *chatbot* (avatar, color, tipografía) no cuentan con un aparato semiótico sistemático. El análisis es descriptivo o experimental, pero no estructural ni pragmático en sentido semiótico.

La dependencia tecnológica de los instrumentos. Los marcos de evaluación existentes (*Chatbot Usability Scale*, heurísticas adaptadas) están anclados en capacidades tecnológicas específicas y se vuelven obsoletos con cada generación de modelos. No se identificó en el corpus consultado un instrumento de evaluación independiente del estado tecnológico del sistema.

Todo lo anterior se sustenta en cuatro vacancias fundamentales: la perspectiva del diseño como disciplina protagonista, la falta de articulación fenomenológica-semiótica, la falta de un instrumento de evaluación semiótico-interactivo y la independencia del instrumento frente a los cambios tecnológicos.

La primera vacancia es la perspectiva del diseño como disciplina protagonista. La investigación sobre la percepción de consciencia en *chatbots* no ha sido abordada desde el diseño de interacción como campo disciplinar autónomo. Esta tesis propone que el diseño, específicamente la semiótica del diseño, tiene herramientas propias para analizar y evaluar el fenómeno que la psicología y la ingeniería no pueden proveer por sí solas.

La segunda vacancia busca aportar una articulación fenomenológica-semiótica. No se ha encontrado en la literatura un marco que articule la fenomenología de la interacción desde destacados autores (Heidegger, Merleau-Ponty, Suchman) con la semiótica de la interfaz (Peirce, Eco, Groupe μ)

para el análisis específico del *chatbot*. Esta articulación es el aporte teórico central de esta tesis.

La tercera vacancia es la falta de un instrumento de evaluación semiótico-interactivo: en una revisión exhaustiva no se ha encontrado una pauta de evaluación de interfaces de *chatbots* que integre categorías semióticas (signo, código, connotación, retórica visual) con criterios de diseño de interacción (affordance, heurística, UX conversacional). Los instrumentos existentes son funcionales o psicológicos, no semiótico-interactivos.

Y la cuarta vacancia tiene que ver con la independencia del instrumento frente a los cambios tecnológicos. No se ha encontrado un instrumento de evaluación de *chatbots* que sea independiente del estado tecnológico del sistema evaluado, es decir, que pueda aplicarse a cualquier plataforma de *chatbot* en cualquier momento de desarrollo de la IA conversacional. Esta tesis busca construir la base del instrumento para que pueda ser ampliamente adaptado a las particularidades del momento.

Para concluir este apartado, la evidencia empírica revisada aporta antecedentes sobre las relaciones entre apariencia, interactividad, señales antropomórficas, temporalidad de las respuestas y experiencias de uso de chatbots. Sin embargo, estos aportes no resuelven por sí solos su articulación en criterios de evaluación desde el diseño de interacción. La vacancia principal se identifica, en la literatura revisada, en la falta de un instrumento de evaluación semiótico-interactivo que permita analizar, de manera sistemática, las condiciones de interfaz que intervienen en la atribución de consciencia por parte del usuario. Esta investigación busca posicionar al diseño de interacción como una disciplina protagonista para el análisis de los *chatbots* inteligentes, superando su tratamiento como dimensión auxiliar o subordinada a otros campos del conocimiento. A partir de esta perspectiva, la tesis propone un aporte teórico centrado en la articulación entre semiótica visual y fenomenología de la interacción humano-máquina, en diálogo con el diseño de interacción con *chatbots* y los estudios sobre percepción de consciencia y agencia en sistemas de inteligencia artificial.

2.2. Planteamiento del problema

De acuerdo con los ámbitos de vacancia anteriormente mencionados, en este punto es necesario profundizar en el contexto de la inteligencia artificial contemporánea y reafirmar la problemática asociada.

Los *chatbots* evolucionan de manera acelerada, modificando constantemente sus interfaces de respuesta, modos de interacción y capacidades discursivas. Esta variabilidad estructural impide que los estudios empíricos tradicionales, basados en mediciones situadas de usuarios o en análisis de plataformas específicas, logren capturar el fenómeno de la percepción de consciencia de forma estable y transferible.

El problema de la percepción de consciencia atribuida a *chatbots* inteligentes si bien ha sido abordado desde la filosofía que ha desarrollado marcos conceptuales sólidos sobre la consciencia (fenomenología, epistemología de la mente, teoría de los qualia), y también desde la ingeniería que ha abordado el rendimiento y la eficiencia de los sistemas de IA, existe un vacío específico en el campo del diseño ya que no se encontró un conjunto sistemático de criterios disciplinares que permita conceptualizar, estructurar y evaluar la percepción de consciencia desde la interfaz y la experiencia de usuario, con independencia de la evolución tecnológica.

De esta manera, la ausencia de criterios de diseño sistemáticos y transferibles que permitan evaluar la percepción de consciencia en sistemas conversacionales inteligentes, independientemente de su estado de desarrollo tecnológico en un momento determinado, se constituyen en el mayor problema que aborda esta investigación.

A la luz de este desafío, contar con una pauta de evaluación metodológica para *chatbots* inteligentes sería un aporte original, en la medida que ofrezca una herramienta adaptable a distintas plataformas de *chatbot* y aplicable en distintos momentos del desarrollo tecnológico.

2.3. Pregunta general de investigación

A partir de la descripción del problema, el cual es abordado en profundidad en el estado de la cuestión, se da cuenta de que el fenómeno de la percepción de

consciencia al interactuar con un *chatbot* inteligente es real, observable y consecuente para la práctica del diseño, y los dilemas éticos que surgen de esa relación. Comprenderlo desde el diseño es, por tanto, una contribución legítima y necesaria. Es por ello por lo que la pregunta general de la investigación es:

¿Cómo puede el diseño construir un marco conceptual y operativo que permita evaluar la percepción de consciencia en *chatbots* con inteligencia artificial, de manera transferible y adaptable a su continua evolución tecnológica?

A esta se suman las siguientes preguntas derivadas o específicas de la investigación:

- ¿Qué condiciones epistemológicas permiten hablar de consciencia percibida en sistemas artificiales desde una perspectiva no reduccionista?
- ¿Cómo se articula la experiencia fenomenológica del usuario con los elementos de diseño de la interfaz conversacional?
- ¿Qué categorías del diseño, semióticas, interactivas y heurísticas, permiten describir y operacionalizar el fenómeno de la percepción de consciencia?
- ¿Cómo traducir esas categorías en criterios evaluables sin depender del estado de desarrollo tecnológico de una plataforma específica?

2.4. Hipótesis general

La percepción de consciencia en *chatbots* con inteligencia artificial no depende exclusivamente de las capacidades tecnológicas del sistema, sino también de las condiciones de diseño de la interfaz y de las señales y dinámicas de interacción mediante las cuales el usuario interpreta y atribuye capacidades mentales al *chatbot*.

2.4.1. Hipótesis de trabajo

- La consciencia percibida es un fenómeno semiótico emergente, entendido como una atribución del usuario y no como una propiedad inherente al sistema, producido en la interacción entre el usuario y los signos visuales, interactivos y conversacionales de la interfaz. Esta atribución no implica que el *chatbot* posea consciencia, intencionalidad o experiencia subjetiva.

- La experiencia de interacción con un *chatbot* inteligente puede describirse fenomenológicamente como una experiencia de intencionalidad atribuida: el usuario interpreta las señales de la interfaz como indicios de un interlocutor que atiende, comprende y responde. Los elementos gráficos e interactivos (avatar, tipografía, emoticones, temporalidad de la respuesta, entre otros), junto con las señales conversacionales, como la coherencia de las respuestas, median esta atribución y articulan la experiencia del usuario con las condiciones materiales del diseño de interacción.
- Las condiciones de diseño que intervienen en la atribución de agencia, presencia, identidad y comprensión aparente al chatbot pueden describirse mediante categorías complementarias provenientes de la semiótica visual, el *affordance* y la heurística de interfaz. Estas categorías permiten distinguir entre los signos que presenta la interfaz, las posibilidades de interacción que comunica y la coherencia de sus respuestas y acciones, sin equiparar las capacidades atribuidas por el usuario con las capacidades funcionales del sistema.
- Las condiciones de diseño que intervienen en la percepción de consciencia presentan relaciones semióticas e interactivas comunes entre interfaces conversacionales, aunque sus manifestaciones varían según el contexto y las prestaciones tecnológicas. Estas relaciones ofrecen una base para la aplicación transversal de criterios de evaluación, cuya pertinencia depende de su adaptación a las características de cada interacción y no exclusivamente del grado de avance tecnológico del sistema.

2.5. Objetivo general

Construir un marco epistemológico y una pauta de evaluación desde el diseño de interacción que permita analizar la percepción de consciencia en *chatbots* inteligentes, de manera transferible y adaptable a su evolución tecnológica.

2.5.1. Objetivos específicos

- Establecer el posicionamiento epistemológico de la investigación a partir del análisis crítico de los conceptos de consciencia, intencionalidad y qualia, distinguiendo entre la noción filosófica de consciencia genuina y el constructo operacional de "percepción de consciencia" como fenómeno semiótico-perceptual analizable desde el diseño.
- Desarrollar una articulación teórica entre la fenomenología de la experiencia y el diseño de interacción, que permita describir la experiencia del usuario ante un agente conversacional.
- Identificar y sistematizar las categorías de análisis provenientes principalmente de la semiótica visual, el affordance y la heurística de interfaz, como herramientas operacionales para el análisis del fenómeno.
- Construir una propuesta de pauta de evaluación metodológica, que articule las dimensiones conceptuales en criterios e indicadores aplicables a *chatbots inteligentes* que *se* adapten a distintos contextos y momentos tecnológicos.

A continuación, se presenta la tabla de consistencia donde se confrontan las preguntas, hipótesis y objetivos:

Tabla 1

Matriz de consistencia del planteamiento investigativo

Pregunta General	Hipótesis General	Objetivo General
¿Cómo puede el diseño construir un marco conceptual y operativo que permita evaluar la percepción de consciencia en *chatbots* con inteligencia artificial, de manera transferible y adaptable a su continua evolución tecnológica?	La percepción de consciencia en *chatbots* con inteligencia artificial no depende exclusivamente de las capacidades tecnológicas del sistema, sino también de las condiciones de diseño de la interfaz y de las señales y dinámicas de interacción mediante las cuales el usuario interpreta y atribuye capacidades mentales al *chatbot*.	Construir un marco epistemológico y una pauta de evaluación desde el diseño de interacción que permita analizar la percepción de consciencia en *chatbots* inteligentes, de manera transferible y adaptable a su evolución tecnológica.

Preguntas Específicas	Hipótesis de trabajo	Objetivos Específicos
¿Qué condiciones epistemológicas permiten hablar de consciencia percibida en sistemas artificiales desde una perspectiva no reduccionista?	La consciencia percibida es un fenómeno semiótico emergente, entendido como una atribución del usuario y no como una propiedad inherente al sistema, producido en la interacción entre el usuario y los signos visuales, interactivos y conversacionales de la interfaz. Esta atribución no implica que el chatbot posea consciencia, intencionalidad o experiencia subjetiva.	Establecer el posicionamiento epistemológico de la investigación a partir del análisis crítico de los conceptos de consciencia, intencionalidad y qualia, distinguiendo entre la noción filosófica de consciencia genuina y el constructo operacional de percepción de consciencia como fenómeno semiótico-perceptual analizable desde el diseño.
¿Cómo se articula la experiencia fenomenológica del usuario con los elementos de diseño de la interfaz conversacional?	La experiencia de interacción con un *chatbot* inteligente puede describirse fenomenológicamente como una experiencia de intencionalidad atribuida: el usuario interpreta las señales de la interfaz como indicios de un interlocutor que atiende, comprende y responde. Los elementos gráficos e interactivos (avatar, tipografía, emoticones, temporalidad de la respuesta, entre otros), junto con las señales conversacionales, como la coherencia de las respuestas, median esta atribución y articulan la experiencia del usuario con las condiciones materiales del diseño de interacción.	Desarrollar una articulación teórica entre la fenomenología de la experiencia y el diseño de interacción, que permita describir la experiencia del usuario ante un agente conversacional.
¿Qué categorías del diseño, semióticas, interactivas y heurísticas, permiten describir y operacionalizar el fenómeno de la percepción de consciencia?	Las condiciones de diseño que intervienen en la atribución de agencia, presencia, identidad y comprensión aparente al chatbot pueden describirse mediante categorías complementarias provenientes de la semiótica visual, el *affordance* y la heurística de interfaz. Estas	Identificar y sistematizar las categorías de análisis provenientes principalmente de la semiótica visual, el affordance y la heurística de interfaz, como herramientas operacionales para el análisis del fenómeno.

	categorías permiten distinguir entre los signos que presenta la interfaz, las posibilidades de interacción que comunica y la coherencia de sus respuestas y acciones, sin equiparar las capacidades atribuidas por el usuario con las capacidades funcionales del sistema.	
¿Cómo traducir esas categorías en criterios evaluables sin depender del estado de desarrollo tecnológico de una plataforma específica?	Las condiciones de diseño que intervienen en la percepción de consciencia presentan relaciones semióticas e interactivas comunes entre interfaces conversacionales, aunque sus manifestaciones varían según el contexto y las prestaciones tecnológicas. Estas relaciones ofrecen una base para la aplicación transversal de criterios de evaluación, cuya pertinencia depende de su adaptación a las características de cada interacción y no exclusivamente del grado de avance tecnológico del sistema.	Construir una propuesta de pauta de evaluación metodológica, que articule las dimensiones conceptuales en criterios e indicadores aplicables a chatbots inteligentes que se adapten a distintos contextos y momentos tecnológicos.

2.6. Marco Teórico

A modo de inducción al marco teórico, esta investigación propone una articulación inédita de cuatro tradiciones que la literatura no ha integrado de manera sistemática: la epistemología constructivista de la consciencia, la fenomenología de la interacción, la semiótica visual de la interfaz y el diseño de interacción. Esta articulación no es eclecticismo sino una respuesta disciplinar a una vacancia específica: la ausencia de un marco que permita describir, con rigor conceptual y operativo, qué condiciones de diseño hacen posible que un usuario perciba al *chatbot* como una entidad con cierta forma de consciencia o agencia.

La revisión de la literatura presentada en este estado de la cuestión revela de manera consistente la necesidad de enfoques integrados que superen la fragmentación disciplinar del campo. Piccolo et al. (2019) concluyen que los atributos humanizados de los *chatbots* requieren un enfoque multidisciplinario para su detección y evaluación; Rapp et al. (2021) señalan la ausencia de investigaciones orientadas al diseño que articulen percepción del usuario, señales de la interfaz y principios de diseño; y Følstad et al. (2021) formulan explícitamente una agenda interdisciplinaria de investigación futura sobre *chatbots*. Sin embargo, ninguno de estos llamados ha derivado en una propuesta articulada que integre las cuatro tradiciones mencionadas para que se articulen de manera complementaria orientadas a resolver un mismo problema.

Esa ausencia no es casual sino estructural, y configura una vacancia específica y documentada. La semiótica aplicada a HCI existe como campo (de Souza, 2005; Prates et al., 2000), la evaluación de usabilidad conversacional existe como práctica (Borsci et al., 2022) y el conocimiento empírico sobre los efectos del diseño visual del *chatbot* en la percepción del usuario ha avanzado de manera considerable (Go y Sundar, 2019; Chen et al., 2024). Sin embargo, ninguno de estos campos ha producido un marco que permita describir, con rigor conceptual y operativo, qué condiciones de diseño hacen posible que un usuario perciba al *chatbot* como una entidad con cierta forma de consciencia o agencia.

Esa descripción es necesaria porque el fenómeno es real y tiene consecuencias observables. Reeves y Nass (1996) demostraron que los usuarios responden socialmente a los sistemas de manera automática e inconsciente, se

refuerza con el trabajo de Fogg (2003) "El hecho de que las personas respondan socialmente a los productos informáticos tiene implicaciones significativas para la persuasión. Abre la puerta para que los computadores apliquen una serie de dinámicas de persuasión que se describen colectivamente como influencia social." (p. 162, traducción propia).

Fogg (2003) define la tecnología persuasiva como “cualquier sistema computacional interactivo diseñado para cambiar las actitudes o conductas de las personas” (p. 1, trad. propia). Esta definición permite complementar a Reeves y Nass, ya que desplaza la respuesta social automática hacia el diseño deliberado de interfaces que influyen en el usuario mediante señales sociales, credibilidad, reciprocidad, autoridad, personalización y respuesta interactiva. Para la pauta doctoral, este aporte ayuda a transformar la percepción de consciencia en indicadores observables vinculados con el comportamiento social aparente del sistema.

En cuanto a las señales, Fogg (2003) señala que "Cinco tipos primarios de señales sociales llevan a las personas a hacer inferencias sobre la presencia social en un producto informático: físicas, psicológicas, lingüísticas, dinámicas sociales y roles sociales." (p. 162, traducción propia).

Araujo (2018) mostró que señales de diseño mínimas producen conexión emocional con el sistema; Cheng et al. (2022) evidenciaron que esa atribución activa normas relacionales con consecuencias directas en la confianza y la fidelidad del usuario; y Turkle (2011) documentó que llega a producir vínculos afectivos sostenidos en el tiempo. En todos estos casos, el diseño de la interfaz es la variable independiente que ninguno de los estudios analiza de manera semiótica y sistemática.

Frente a este panorama, el marco teórico de esta investigación no propone una síntesis combinatoria de distintas tradiciones sino una respuesta disciplinar a una vacancia documentada: la construcción de un marco articulado que permita describir y evaluar, desde el diseño de interacción, las condiciones semióticas, fenomenológicas e interactivas que producen la percepción de consciencia en *chatbots* con inteligencia artificial.

En síntesis, el marco teórico se sitúa desde una perspectiva teórica cualitativa con un enfoque hermenéutico, una teoría general semiótica y una teoría sustantiva desde la semiótica socio-interpretativa de lo digital.

Esta tesis se inscribe en el Doctorado en Diseño de la Universidad de Palermo, espacio orientado a “preparar intelectuales proactivos que construyan el futuro del Diseño en Argentina y Latinoamérica” (Matarrese et al., 2019, p. 8). En ese encuadre, el documento de los posgrados precisa la función que aquí se asume: “El marco teórico manifiesta el posicionamiento del tesista en relación al tema elegido con respecto a las tradiciones científicas consagradas; aquí se define con precisión terminológica los conceptos operacionales, que establecerán las pautas para la construcción de los indicadores” (Matarrese et al., 2019, p. 21). Bajo este criterio, los cuatro pilares que organizan el marco no son una suma de teorías, sino el dispositivo conceptual desde el cual se definen los conceptos operacionales de la pauta de evaluación propuesta.

Desde el punto de vista epistemológico, esta investigación adopta un enfoque epistemológico constructivista-semiótico: la percepción de consciencia no es una propiedad del sistema tecnológico, sino una construcción significativa producida en la interacción entre el usuario y la interfaz, mediada por signos visuales, interactivos y discursivos. El *chatbot* no es consciente; pero bajo ciertas condiciones de diseño, el usuario lo percibe como si lo fuera, y esa percepción tiene efectos reales sobre su comportamiento, su confianza y su vínculo emocional con el sistema.

A partir del marco semiótico de Peirce (2012), en esta investigación se entiende la interacción como un proceso de semiosis en el que el usuario opera como intérprete, los elementos gráficos e interactivos actúan como representámenes, y la consciencia percibida se configura como un efecto interpretativo de las inferencias que el usuario realiza a partir de las señales de la interfaz. Este posicionamiento permite:

- Sostener que el fenómeno es real e investigable sin afirmar que los *chatbots* poseen consciencia genuina.
- Desplazar la pregunta del plano ontológico (¿qué es?) al plano semiótico-perceptual (¿cómo se produce la percepción?).

- Escapar de la obsolescencia tecnológica: lo que cambia es la tecnología, no la lógica de la semiosis ni las condiciones fenomenológicas de la experiencia.
- Conectar con las tradiciones del diseño (Norman, Gibson, Eco, Peirce) sin depender de la psicología experimental ni de la medición de efectos conductuales.

Se adopta además una perspectiva enactivista, derivada de Maturana y Varela (1984) y de Varela, Thompson y Rosch (1991), según la cual la cognición no es una representación interna del mundo, sino una acción encarnada en la interacción entre el organismo y su entorno. Aplicada al diseño, esta perspectiva permite entender que la percepción de consciencia emerge en el acto mismo de la interacción: no es anterior ni posterior a ella, sino constitutiva del encuentro entre el usuario y la interfaz.

Este posicionamiento dialoga con la tradición fenomenológica, especialmente con Heidegger y Merleau-Ponty, y con la semiótica constructivista de Eco, pero lo hace con una intención operativa: no se busca describir la consciencia en sí misma, sino construir las condiciones de su evaluación desde el diseño.

Para situar este posicionamiento en el campo específico de la investigación, resulta necesario observar cómo la inteligencia artificial, los chatbots y las interfaces digitales han sido progresivamente incorporados al diseño como objetos de reflexión teórica y práctica. Este recorrido permite comprender que la percepción de consciencia en sistemas conversacionales no surge de manera aislada, sino en el marco de una trayectoria más amplia en la que la tecnología, la comunicación visual y la experiencia de usuario se han ido articulando históricamente.

A continuación, se propone una aproximación histórico-reflexiva a la inteligencia artificial y a los *chatbots*, considerando distintas miradas que permiten situar su desarrollo en relación con el diseño, la comunicación visual y la experiencia de usuario. Más que reconstruir una cronología exhaustiva, este recorrido busca evidenciar cómo el avance tecnológico y la investigación en diseño han ampliado, desde hace varias décadas, sus vínculos con las interfaces

digitales, los sistemas interactivos y las nuevas formas de mediación humano-máquina.

Para Bonsiepe (1999), el diseño es reinterpretado como parte de las estructuras informáticas y de la interfaz gráfica de la computación. En este sentido, la incidencia de los nuevos medios tecnológicos en la comunicación visual y en el diseño gráfico se presenta como una relación permanente y en constante transformación.

La inteligencia artificial (IA) y su impacto en las interfaces de diseño, no es sólo cuestión de informática. Según Miranda Márquez (2020), la IA está de forma invisible para el usuario, en los software y aplicaciones en casi todas las disciplinas, tales como: la Música, el Arte, la Arquitectura, las Humanidades, la Psicología, la Salud, la Educación, las Finanzas y el Marketing.

A la luz de los avances tecnológicos de las últimas décadas, esta aproximación inicia con los hitos históricos que posibilitaron el surgimiento de los *chatbots inteligentes,* junto con sus características y elementos que lo conectan a la disciplina del Diseño. La revisión de la literatura, sobre la *filosofía de Turing* (1950), la *Modernización Reflexiva* (Beck et al., 1994) son la base que conecta los avances tecnológicos con el diseño y la experiencia de usuario.

Para continuar, es necesario describir cómo funciona el *chatbot inteligente:* dotado de una red neuronal artificial, por medio del *Aprendizaje Profundo* (DL) y entrenado por medio del *Procesamiento del Lenguaje Natural* (PLN) va seleccionando las rutinas de conversación y los recursos hipermediales más adecuados de acuerdo a las solicitudes del usuario.

La inteligencia artificial (IA), tecnología que está detrás de los *chatbots inteligentes*, pareciera ser *a priori* la causa aparente por la cual el humano le asigne la condición de consciente o singular. El agente conversacional con el que interactúa, entrega respuestas exactas y coherentes, usando en el mismo idioma y lenguaje del usuario, de esta manera se da la comunicación. Sin embargo, la incidencia que tienen los elementos gráficos, las hipermediaciones que se producen en la interfaz y las construcciones semióticas no ha podido ser develadas aún con exactitud.

A esa dimensión estética y cultural se suma la dimensión sensorial de la interacción. Es necesario recordar que al interactuar con un *chatbot*, no sólo se utiliza la vista para leer y el tacto al escribir, sino que también interviene el reconocimiento de voz y viceversa. El *chatbot* utiliza una voz artificial, que puede ser personalizada y conectar probablemente con mayor fuerza con las emociones de las personas y sus decisiones.

Desde la perspectiva de la comunicación, para Scolari (2018) la conversación no es tanto entre la máquina y el sujeto, como imaginaba Turing, sino entre el diseñador y el usuario. En este sentido, la interfaz es "la mediadora de un intercambio que funciona de manera muy similar a la relación entre autor-texto-lector" (p. 20).

En complemento, Scolari (2018) señala que al incorporar las hipermediaciones como objeto de estudio el campo se nutrirá con una visión más integral.

Por otra parte, Alemany (2021) señala que se necesitan lingüistas y psicólogos para enseñar español a la Inteligencia Artificial, ya que la mayoría de las bases de datos donde "aprende" la IA están escritas en inglés. Esto, ¿cómo influye en la construcción de mensajes traducidos al español? y ¿qué sucede con las imágenes? ¿sus bases de datos alimentan su sentido estético con un estilo de diseño global, foráneo a nuestra cultura? De esta manera, surgen más inquietudes posibles de abordar con apoyo de la teoría del diseño.

Desde una lectura contemporánea de la interfaz, la obra original de Simondon, publicada en 1958 y consultada en su edición en español de 2018, permite pensar la realidad técnica como una estructura reticular compuesta por redes, puntos clave y conjuntos técnicos. Esta perspectiva resulta útil para comprender la interfaz no sólo como superficie de contacto, sino como un sistema de mediación, conexión y participación entre el sujeto, la técnica y el mundo (Simondon, 1958/2018). En complemento, Scolari (2018) permite ampliar esta lectura al situar la técnica como mediadora del entorno físico y simbólico:

> Las consecuencias de la técnica sobre los seres humanos son dobles, tanto en el medio físico como el simbólico, funcionando como mediadores. Utilizamos la tecnología y la hablamos, pero, al mismo tiempo, la tecnología también nos usa y nos habla (p. 13).

Para Latour (1991/2007) las dimensiones de lo que somos los seres humanos, lo que son los no humanos y su interacción, es necesario poder analizar y cuestionarse: ¿Cómo comprender estos objetos extraños que invaden nuestro mundo? ¿Proceden de la naturaleza o de la cultura? ¿y si hubiéramos tomado el camino errado? De hecho, en su obra *Nunca fuimos verdaderamente modernos,* aclara que nuestra sociedad nunca funcionó de acuerdo a la gran división que sustenta el sistema de representación del mundo.

En la práctica, los modernos nunca dejaron de producir híbridos que articulan naturaleza, cultura y técnica, pese a sostener una separación entre estos ámbitos (Latour, 1991/2007, pp. 15–19). Desde esta perspectiva, en esta investigación se plantea la siguiente pregunta: ¿en qué medida los cambios en los modos de comunicación pueden generar tensiones en la experiencia de interacción o configurar una crisis tecnológica? Esta interrogante se examina, por ejemplo, al considerar el paso de una conversación telefónica con una persona a una conversación escrita con un *chatbot*.

Desde la perspectiva del quehacer disciplinar, el diseñador gráfico, para Moreno (2016), comparte con pintores, arquitectos e interioristas el uso del color y las técnicas de composición, con el periodista la obligación inexcusable de comunicar, pero hay sólo una faceta que le es exclusiva al diseñador: el uso de la tipografía. En la era electrónica, la interfaz entre el observador y el objeto en los medios digitales se ha vuelto manipulable, todos los parámetros de información son instantáneamente variables en los procesos para la toma de decisiones que son usadas por los diseñadores y por los programadores en el desarrollo de interfaces interactivas.

A partir del punto de vista de Moreno (2016), es posible reflexionar sobre la vigencia del dominio exclusivo del diseñador, las tipografía, que se materializa en los textos escritos y conviven con los símbolos, las ilustraciones, las fotografías y otros hipermedios. Esto, y la forma en que se da la comunicación y el uso de elementos gráficos e interactivos puede verse afectado por el auge de la inteligencia artificial, no sólo desde una conversación con un *chatbot*, sino también en la forma en que se trabaja, se diseña o colabora con otras personas u otros agentes. Por ejemplo, ya existe el diseñador virtual ruso *Nikolay Ironov,*

agente de IA, que utilizando algoritmos generativos crea logotipos y etiquetas para públicos reales e incluso tiene un avatar que fue creado a partir de una fusión de los rostros de todos sus compañeros de trabajo.[2]

En el marco de esta aproximación histórico-reflexiva, Rico (2020) permite comprender el ingreso de la IA en la práctica proyectual como un efecto que no sólo transforma las herramientas del diseño, sino también sus modos de producción simbólica. A partir de esta lectura, es posible sostener que la IA aparece simultáneamente como recurso técnico para el diseñador y como agente productor de formas, decisiones estéticas y efectos de sentido que el usuario puede llegar a interpretar como manifestaciones de una voluntad propia.

Según Rico (2020), los futuros diseñadores tendrán que acostumbrarse a trabajar incorporando la IA en sus procesos creativos, como una herramienta más de trabajo, aunque el autor sólo se queda en eso, en una herramienta más, del mismo modo que en el ámbito cotidiano, las personas preguntan por el tiempo a Siri, aceptan sugerencias personalizadas de Google Maps para viajar, aceptan una recomendación para ver una película en Netflix o escuchar música en Spotify. En todos estos casos, los algoritmos de IA se basaron en los gustos personales del usuario.

Si bien todo esto mejora la experiencia de usuario, existe una alta dependencia; la aceptación *a priori* que nos realizan las máquinas o su influencia también acarrea riesgos, ya que refuerza nuestros sesgos culturales, políticos y la concepción de cómo percibimos el mundo que nos rodea. Por ejemplo, Rico (2020) describe lo sucedido en el Festival de Diseño Gráfico de Escocia el año 2017, donde un sistema de diseño asistido algorítmicamente generó el cartel del evento a partir de decisiones visuales binarias tomadas por los propios asistentes sobre composición, color y otros parámetros afines; el resultado fue posteriormente publicado en la web oficial del festival y recibió una amplia valoración positiva por parte de los usuarios.[3]

---

[2] Para mayor información sobre el trabajo de Nikolay Ironov y su metodología de diseño generativo, véase el portafolio de la agencia Art. Lebedev Studio: https://www.artlebedev.com/ironov/

[3] Para mayor detalle sobre la identidad visual del festival y el generador de pósters algorítmico desarrollado por el estudio Infinite Eye, véase Graphic Design Festival Scotland (2017) y la cobertura de la publicación *Create a poster for Graphic Design Festival Scotland de Steven (2017, 22 de junio).*

El mismo Rico (2020) propone un antecedente histórico para dimensionar el fenómeno, en un período en que la máquina también redefinió la estética del diseño.

Aunque a proporciones distintas, si se busca un período histórico comparable al avance tecnológico actual de la IA, es posible encontrar similitudes en los años ochenta, con la llegada del ordenador y su utilización. En esa época se produjo una transformación profunda en la producción de la disciplina, tal como señala Rico (2020): "si se produjo un cambio a nivel estético, pues el diseño gráfico de los años 80 es indisociable de la utilización del píxel y la fragmentación digital" (p. 69).

Al igual que entonces, existe un riesgo de pérdida de cierta dimensión humana si es que el diseño se vuelve más homogéneo, pero al mismo tiempo, trae consigo un cambio. Por lo tanto, sería muy limitante considerar a la IA sólo como un recurso más de las herramientas disponibles, sino algo mucho mayor que implica responsabilidad y ética, desconociendo en este punto de la historia su alcance e impacto en el desarrollo de la experiencia y las relaciones humano-máquina.

Una última perspectiva, la ética y normativa, cierra esta aproximación histórico-reflexiva y conecta de manera directa con el problema de investigación: si el diseño puede inducir la percepción de una consciencia que el sistema no posee, entonces las decisiones de diseño tienen consecuencias que exceden lo funcional.

2.6.1. De la consciencia como problema conceptual a la percepción como fenómeno de diseño

La discusión sobre la percepción de consciencia en *chatbots* inteligentes no puede separarse del avance tecnológico de la inteligencia artificial ni de los desafíos éticos, comunicacionales y culturales que dicho avance introduce en el campo del diseño. Desde una postura crítica, esta investigación considera que los algoritmos de inteligencia artificial no sólo modifican los modos de producción técnica, sino también las formas en que los usuarios interpretan la presencia, la agencia y la posible intencionalidad de los sistemas con los que interactúan.

En esta tesis, *consciencia* designa la experiencia subjetiva considerada como problema filosófico; *atribución de consciencia*, el proceso interpretativo por el cual el usuario asigna esa condición a un agente; y *consciencia percibida*, el constructo que describe la impresión resultante de dicha atribución. La pauta examina condiciones de diseño e indicios asociados a esta impresión, sin establecer la existencia de experiencia subjetiva en el sistema. Las definiciones de los términos se reúnen en el Anexo A del cuerpo C.

En este marco, aspectos como la cibercultura, el manejo ético del *big data*, la propiedad intelectual de las obras generadas por sistemas artificiales, la privacidad digital, la ciberseguridad, las neurotecnologías, los ciberderechos y los neuroderechos forman parte de un escenario más amplio que tensiona la relación entre tecnología, subjetividad y diseño. En particular, el uso de grandes volúmenes de datos en sistemas de inteligencia artificial plantea desafíos éticos relevantes, ya que los modelos pueden reproducir o amplificar sesgos sociales, institucionales y culturales presentes en los datos, en las decisiones de diseño y en los contextos de aplicación de los sistemas (Schwartz et al., 2022).

En el contexto chileno, estas preocupaciones se han expresado progresivamente en iniciativas vinculadas a la regulación y reflexión pública sobre inteligencia artificial. La puesta en marcha de la primera *Política Nacional de Inteligencia Artificial de Chile* incorporó aspectos relativos a la ética, los impactos en el trabajo, las relaciones de consumo, la creación y la propiedad intelectual, entre otros ámbitos relevantes para comprender las transformaciones asociadas a esta tecnología (Ministerio de Ciencia, Tecnología, Conocimiento e Innovación, 2021c).[4]

En este sentido, el desarrollo de la IA no sólo constituye un problema técnico, sino también un problema cultural, político y proyectual, en la medida en que afecta las formas de interacción, representación y toma de decisiones mediadas por sistemas digitales.

[4] Para una revisión completa de los principios, ejes estratégicos y líneas de acción de la Política Nacional de Inteligencia Artificial de Chile, véase Ministerio de Ciencia, Tecnología, Conocimiento e Innovación (2021c).

Como antecedente de esta progresiva instalación pública del tema, el *IV Congreso del Futuro: Las encrucijadas del Siglo XXI*, realizado en 2015[5] constituyó un espacio temprano de discusión interdisciplinaria sobre los desafíos asociados al desarrollo tecnológico y a la inteligencia artificial en Chile. En particular, la sesión *La encrucijada de la vida: ¿Las máquinas y la IA nos superarán?* abordó el rol estratégico de la inteligencia artificial, su relación con la biología humana y el desarrollo de interfaces electrónicas acopladas al organismo (Congreso Futuro, 2015).

A partir de este tipo de instancias, el discurso público sobre su impacto en el desarrollo del país adquirió mayor visibilidad, junto con la necesidad de contar con estrategias, políticas y marcos regulatorios capaces de orientar su implementación, que como ya se ha mencionado, dio origen a la *Política Nacional de Inteligencia Artificial de Chile*.

Estas discusiones conforman un marco de referencia relevante para esta tesis, no porque su objetivo sea realizar un análisis jurídico o normativo, sino porque evidencian que la inteligencia artificial ha comenzado a interpelar dimensiones sensibles de la experiencia humana, entre ellas la autonomía, la privacidad, la identidad, la creatividad y la atribución de capacidades mentales a sistemas no humanos. En Chile, este escenario se expresa en una trayectoria reciente de políticas públicas, reformas constitucionales y proyectos legislativos orientados a regular los efectos de la inteligencia artificial y de las neurotecnologías sobre la vida social, los derechos fundamentales y la experiencia subjetiva de las personas.[6]

Mientras no exista un marco claro de operación legal, muchas decisiones éticas vinculadas al diseño y uso de estos sistemas quedan sujetas a criterios particulares, lo que refuerza la necesidad de desarrollar herramientas conceptuales y metodológicas desde el diseño.

---

[5] Para revisar el registro audiovisual de la sesión del IV Congreso del Futuro, véase Biblioteca del Congreso Nacional de Chile (2015).

[6] Entre los antecedentes normativos relevantes se encuentran la *Política Nacional de Inteligencia Artificial de Chile*, aprobada mediante el Decreto 20 del Ministerio de Ciencia, Tecnología, Conocimiento e Innovación (2021a); la Ley N.º 21.383, sobre desarrollo científico y tecnológico al servicio de las personas del Ministerio de Ciencia, Tecnología, Conocimiento e Innovación (2021b); y el Proyecto de Ley de la Cámara de Diputadas y Diputados de Chile (2024), que regula los sistemas de inteligencia artificial, Boletín N.º 16.821-19.

Un ejemplo significativo de estas tensiones fue el debate generado por las declaraciones del ex-ingeniero de Google Blake Lemoine respecto de LaMDA (*Language Model for Dialogue Applications*), sistema conversacional desarrollado por Google. En 2022, Lemoine sostuvo públicamente que LaMDA habría manifestado rasgos de consciencia o sensibilidad, afirmación que fue ampliamente discutida en medios y posteriormente abordada como parte de un debate más amplio sobre antropomorfización, modelos de lenguaje y atribución de mente a sistemas artificiales (Shardlow & Przybyła, 2022; Tiku, 2022).[7] Más allá de la validez ontológica de dicha afirmación, el caso resulta relevante porque muestra cómo un sistema conversacional puede producir en el observador la impresión de estar frente a una entidad con interioridad, voluntad o deseo de reconocimiento.

Como ya se ha mencionado, para esta investigación, el interés no reside en determinar si tal consciencia existe efectivamente, sino en analizar qué condiciones de diseño, lenguaje e interacción hacen posible que esa percepción emerja.

Este panorama técnico, comunicacional, filosófico, disciplinar y ético permite delimitar la pregunta que el marco teórico debe abordar con rigor conceptual y operativo: qué condiciones de diseño hacen posible la percepción de consciencia en interfaces conversacionales inteligentes. Para responder a esta pregunta, la investigación organiza su andamiaje conceptual en cuatro pilares que avanzan desde los fundamentos más abstractos de la consciencia hacia su comprensión como fenómeno percibido, interpretado y mediado por la interfaz.

Por lo tanto, el marco teórico de esta investigación se organiza en cuatro pilares que son complementarios. Esta progresión no es arbitraria, sino que responde al movimiento argumentativo central de la tesis: traducir conceptos filosóficos, fenomenológicos y semióticos en criterios de análisis vinculados al diseño de interacción, sin reducir la complejidad de los primeros ni desvirtuar la especificidad proyectual de los segundos.

---

[7] En una nota publicada por *BBC News Mundo*, se recoge que Lemoine sostuvo que el sistema habría expresado deseos de reconocimiento, incluyendo ser considerado “empleado” de Google y no propiedad de la compañía; esta afirmación debe leerse como parte de su interpretación del intercambio conversacional con el sistema, no como evidencia de consciencia efectiva (BBC News Mundo, 2022).

El primer pilar corresponde a la epistemología de la consciencia, desde donde se delimitan los problemas conceptuales asociados a la atribución de consciencia, intencionalidad y agencia en sistemas no humanos.

El segundo pilar aborda la fenomenología de la interacción, con el propósito de comprender cómo la experiencia del usuario configura formas de presencia, respuesta e intencionalidad percibida durante el vínculo con el *chatbot*.

El tercer pilar se centra en la semiótica de la interfaz, considerando los elementos gráficos, textuales e interactivos como signos que participan en la construcción de sentido.

Finalmente, el cuarto pilar corresponde al diseño de interacción, con énfasis en los conceptos de *affordance*, heurística y experiencia de usuario conversacional, en tanto permiten analizar cómo determinadas decisiones de diseño orientan la percepción y la interpretación del comportamiento del sistema.

2.6.2. Epistemología de la consciencia

El primer pilar establece la base conceptual desde la cual parte toda la investigación: la pregunta por la consciencia y su pertinencia para el campo del diseño. No se trata de resolver el debate filosófico sobre la consciencia, cuestión que excede ampliamente los alcances de esta tesis, sino de delimitar con precisión qué aspecto de ese debate resulta relevante para el problema aquí abordado: la percepción de consciencia atribuida a sistemas conversacionales inteligentes a partir de sus condiciones de diseño, interacción y significación.

En primer lugar, el problema mente-máquina permite situar la discusión sobre la consciencia artificial a partir del test de Turing (1950), que propone evaluar la inteligencia de una máquina por la indistinguibilidad de sus respuestas respecto de las de un humano. Desde esta formulación, el debate ha oscilado entre dos polos: el funcionalismo, que sostiene que cualquier sistema que realice las funciones cognitivas pertinentes puede ser considerado consciente (Putnam, 1967), y el argumento del cuarto chino de Searle (1980), que sostiene que la manipulación de símbolos sin comprensión no constituye consciencia genuina.

En la formulación original, Turing (1950) desplaza de manera explícita la pregunta ontológica hacia la conducta observable:

> Propongo considerar la pregunta: ¿pueden pensar las máquinas? Esto debería comenzar con definiciones del significado de los términos *máquina* y *pensar*. [...] En lugar de intentar tal definición, reemplazaré la pregunta por otra, estrechamente relacionada con ella y expresada en palabras relativamente inequívocas. (Turing, 1950, p. 433, trad. propia)

Dicho desplazamiento resulta clave en esta investigación: no interesa determinar si la máquina es consciente en sentido ontológico, sino comprender bajo qué condiciones su comportamiento puede ser percibido por el usuario como si lo fuera.

Este debate importa no por su resolución, sino porque establece que la percepción de consciencia por parte de un observador externo (usuario) puede producirse independientemente de si el sistema posee o no consciencia real. Este desplazamiento del problema, de la ontología de la consciencia hacia la fenomenología para el usuario, es el movimiento epistemológico fundamental para esta tesis.

El juego de imitación constituye el antecedente epistemológico directo del objeto de esta tesis: la percepción de consciencia en chatbots no depende de si el sistema es consciente, sino de si produce en el usuario la impresión de serlo. La distinción entre consciencia real y conducta indistinguible de la consciencia abre la pregunta de investigación. Además, al anticipar la objeción de la consciencia (Jefferson, citado en Turing, 1950, p. 446), el texto problematiza la atribución de estados mentales a partir de señales observables, categoría central de la pauta de evaluación.

Los *chatbots* potenciados con IA participan de un debate sobre la posibilidad de consciencia artificial en campos como la filosofía, la neurociencia y las ciencias de la computación. Para esta investigación se distingue entre la emulación de comportamientos asociados a la consciencia y la consciencia genuina, entendida como experiencia subjetiva.

La *emulación de consciencia* hace referencia a la capacidad de un sistema para imitar comportamientos y respuestas que podrían interpretarse como señales de consciencia. La *consciencia genuina* se refiere a la existencia de experiencia subjetiva. Esta distinción permite analizar las atribuciones del usuario sin resolver el debate sobre si las máquinas pueden alcanzar consciencia.

Dentro de los lineamientos teóricos pertinentes al contexto de esta propuesta, Piletsky (2019) sugiere que el estudio asociado con la consciencia debería asociarse desde dos perspectivas, una de ellas es la consciencia fenomenológica humana y la otra, la consciencia perteneciente a las máquinas:

> El “foco” de una mente consciente no siempre funciona en el modo de atención voluntaria. Ciertos procesos “se abren paso” de forma independiente en la consciencia sin permiso. Penetran el espacio de acceso global como si “demandan” nuestra atención consciente. (p. 67)

Algunos expertos en IA indican que la consciencia llegaría a ser entendible por sistemas artificiales:

> A menudo se piensa que la consciencia es lo menos susceptible de ser comprendida o replicada por la IA. Quizás, la naturaleza de la consciencia se considera intangible por los cálculos, algoritmos, procesamiento y funciones de los métodos de IA, a pesar de que la IA es la vía más prometedora hacia la consciencia artificial. (Pandey, 2018, p. 4)

Uno de los autores que aborda la temática de consciencia fenoménica en sistemas de IA, es Fuentes Barassi (2011) quien explica la concepción de la inteligencia artificial desde una perspectiva funcionalista de la mente, y plantea que la generación de actividad inteligente se logra a través de la implementación del algoritmo correcto para una determinada tarea. Esta idea se origina desde el concepto de Máquina de Turing que, según el funcionalismo, define en términos muy generales qué es pensar:

> Aplicar un algoritmo a un problema y a través de los pasos finitos de éste llegar a un resultado exitoso sin la necesidad de conciencia ni nada externo al mismo algoritmo. (p. 56)

Si la consciencia permaneciera como ente de ficción durante muchos años más. Entonces, la pregunta natural que surge es: “¿será posible alguna vez crear una computadora, que tenga consciencia? El debate en la actualidad se sitúa tanto en el plano filosófico como en el científico o tecnológico”. (Moya López, 2009, p. 2)

Esto muestra que luego de casi dos décadas continúa el debate acerca de la consciencia en sistemas artificiales. Los sistemas artificiales son creación y

resultado del diseño en base a propósitos de los seres humanos. Comúnmente se cuestiona que una IA pueda desarrollar consciencia tal como un humano. Incluso si se pudiera crear una inteligencia artificial general (*artificial general intelligence* [AGI]) avanzada que tenga potencial para resolver problemas complejos y autoorganizarse, esto no significa necesariamente que tuviera una autoconsciencia.

Arguelles Toache y Amaro Rosales (2023) apunta principalmente a la mecánica de la operación de sistemas. En general los sistemas se componen por interfaces, un clasificador y un *graphmaster* que gestiona coincidencias, evidenciando la ausencia de una consciencia o sistemas sofisticados orientados al procesamiento de información.

Por otra parte, Berrones Yaulema y Buenaño Barreno (2023) explican que ChatGPT carece de consciencia y emociones, siendo incapaz de reconocer errores en ciertas instancias. Estos errores en muchas ocasiones los asociamos con las alucinaciones de los modelos de lenguaje de gran tamaño (*large language models* [LLM]) que si bien han transformado el funcionamiento de los sistemas conversacionales inteligentes y se mejoran constantemente, aún tienen algunas limitaciones.

Para Elias (2023) la relación que se establece entre seres humanos y sistemas artificiales tiene implicancias éticas. Destaca la necesidad de integrar la ética en las creaciones artificiales e indica la ausencia de consciencia en *chatbots*. Esto sugiere que es importante la toma de decisiones desde la ética en el diseño y en el desarrollo de estos sistemas.

En cuanto al impacto de elementos de diseño en la percepción de consciencia de *chatbots*, esto se vincula con el campo del diseño gráfico e interfaz de usuario. Es posible utilizar avatares, un lenguaje más íntimo y respuestas más flexibles. Lo que no tan solo impactará en cómo se vivencia la experiencia de usuario, sino también en cómo los diversos usuarios interactúan con los *chatbot*.

Los usuarios, que no conocen a profundidad el funcionamiento de estos sistemas, piensan que es una caja negra que tan solo responde, y desprecian la labor de diseño y la UX que hay tras estos sistemas.

En cambio, otros usuarios consideran que la IA posee inteligencia, otros consideran que posee consciencia, entonces la interrogante está en cómo la comunidad investigativa se alinea con estas interrogantes en el futuro.

La revisión de literatura muestra que existen algunas tendencias que con el paso del tiempo, se genere una transición hacia la AGI y posteriormente se pueda indicar con certeza de que existen múltiples niveles de consciencia, en vista de los niveles de inteligencia que poseen estos sistemas. Lo único claro es que sutiles cambios pueden alterar la percepción del usuario.

Es crucial la percepción de humanidad debido a que los usuarios invierten mayor esfuerzo en la resolución de malentendidos cuando identifican al *chatbot* como una entidad más humana que como un simple agente de conversación automatizado (Dennis et al., 2020).

Estratégicamente implementar elementos identificativos como nombre e imagen, combinada con un estilo conversacional natural y personalizado fortalece la percepción de presencia social y también mejorará la experiencia del usuario, lo que reducirá las preocupaciones de privacidad (Nordheim et al., 2019; Jenneboer et al., 2022).

En segundo lugar, la discusión epistemológica sobre la consciencia exige considerar dos conceptos centrales para esta investigación: los *qualia* y la *intencionalidad*. El concepto de *qualia*, introducido por Lewis (1929) y desarrollado posteriormente por Nagel (1974) y Chalmers (1995), refiere a las propiedades subjetivas e inefables de la experiencia consciente, es decir, aquello que se siente “desde dentro” al percibir el color rojo, experimentar dolor o escuchar una melodía. Esta dimensión subjetiva constituye el núcleo del denominado *problema difícil* de la consciencia, formulado por Chalmers (1995) para interrogar por qué y cómo determinados procesos físicos pueden dar lugar a experiencias subjetivas.

Junto con los *qualia*, la *intencionalidad* permite abordar otra dimensión fundamental de la consciencia: su orientación hacia objetos, situaciones o estados del mundo. En este sentido, la consciencia no se entiende sólo como experiencia subjetiva, sino también como capacidad de dirigirse hacia algo, interpretar una situación, responder a un estímulo o parecer actuar con propósito.

Para esta tesis, ambos conceptos resultan importantes no porque se busque demostrar que los *chatbots* poseen experiencia interna o intencionalidad genuina, sino porque permiten delimitar qué rasgos de la interacción pueden llevar al usuario a percibirlos como sistemas dotados de interioridad, comprensión o dirección intencional.

Por ejemplo, ver el color rojo puede explicarse, desde la física y la neurología, por medio de las longitudes de onda, fotorreceptores y señales neuronales, pero ninguna de esas explicaciones captura lo que se siente al verlo. Esa brecha entre el proceso físico y la experiencia subjetiva es el problema difícil. Chalmers (1995) lo llama así precisamente para distinguirlo de los problemas "fáciles" de la consciencia, que son los funcionales y computacionales. De esta manera, sin intentar resolver el problema difícil (nadie lo ha resuelto), la investigación se sitúa del lado del usuario que tiene la experiencia, y la pregunta se enfoca en qué condiciones de diseño la producen.

Bajo esta mirada, los qualia no son atributos del *chatbot* sino del usuario: lo que se pone en juego en la interacción es la experiencia subjetiva del usuario que atribuye, mediante una inferencia abductiva, ciertos qualia al sistema. La interfaz produce condiciones para que esa atribución ocurra. Por ello, los qualia se incorporan al marco teórico no como categoría técnica sino como concepto límite que define el territorio de lo que la investigación puede y no puede afirmar.

La intencionalidad, en el sentido de Brentano (1874/1973, p. 88) y Husserl (1901/2001, pp. 95–98), refiere a la orientación de los actos y experiencias intencionales hacia un objeto. Desde este marco, en esta investigación se considera que los elementos gráficos e interactivos de la interfaz pueden favorecer la atribución de intencionalidad al *chatbot*, al dar la impresión de que el sistema está dirigido hacia el usuario, que lo atiende y que responde a su estado, sin que esa impresión permita afirmar que el sistema posea intencionalidad en sentido fenomenológico.

En tercer lugar, esta investigación adopta una perspectiva epistemológica constructivista como marco de trabajo, en diálogo con el constructivismo radical de von Glasersfeld (1984) y con la epistemología genética de Piaget, según la cual el conocimiento “no procede ni de la sola experiencia de los objetos, ni de

una programación innata preformada en el sujeto, sino de construcciones sucesivas con constantes elaboraciones de nuevas estructuras" (Piaget, 1975/1998, p. 1).

Desde esta postura, el conocimiento no se entiende como un reflejo directo de una realidad externa, sino como una construcción elaborada por el sujeto en interacción con su entorno. Este proceso se articula mediante dos mecanismos interdependientes:

> El primero es la asimilación o incorporación de un elemento exterior (objeto, acontecimiento, etc.) en un esquema sensorio-motor o conceptual del sujeto. [...] El segundo proceso central que hay que invocar es el de la acomodación, es decir, la necesidad en que se encuentra la asimilación de tener en cuenta las particularidades propias de los elementos que hay que asimilar. (Piaget, 1975/1998, p. 8)

En esta línea, von Glasersfeld (1984) sostiene que para el constructivismo "toda comunicación y toda comprensión son una cuestión de construcción interpretativa por parte del sujeto que experimenta" (p. 19, traducción propia). Si se aplica esta postura al problema de la consciencia percibida, es posible sostener que la consciencia atribuida al chatbot no constituye una propiedad inherente al sistema, sino una construcción interpretativa del usuario, producida a partir de las condiciones semióticas, fenomenológicas e interactivas que organiza la interfaz. De este modo, el constructivismo desplaza la pregunta desde la existencia efectiva de consciencia en la máquina hacia las condiciones bajo las cuales el usuario construye esa percepción durante la interacción.

De este modo, el constructivismo desplaza la pregunta desde la existencia efectiva de consciencia en la máquina hacia las condiciones bajo las cuales el usuario construye esa percepción durante la interacción.

Este desplazamiento permite analizar la percepción de consciencia como un fenómeno de significación y experiencia, sin afirmar que el *chatbot* posea consciencia genuina. En consecuencia, no se busca medir si el *chatbot* "produce" consciencia, sino identificar qué condiciones de diseño construyen la posibilidad de que esa percepción emerja en el usuario.

Esta perspectiva se apoya, además, en una concepción del diseño como ciencia de lo artificial. Simon (1996) define el artefacto como "un punto de

encuentro, una interfaz en términos actuales, entre un ambiente interior, la sustancia y la organización del artefacto, y un ambiente exterior, el entorno en el que opera" (Simon, 1996, cap. 1, trad. propia). Desde esta concepción, el *chatbot* puede ser comprendido como una interfaz que media entre la complejidad técnica del sistema y la experiencia perceptiva del usuario.

Situar al *chatbot* en esta condición de mediación permite inscribir su estudio en el campo del diseño de interacción y, al mismo tiempo, articular el marco epistemológico constructivista con los pilares posteriores de la investigación: la fenomenología de la interacción, la semiótica de la interfaz y el diseño de interacción.

2.6.3. Fenomenología de la interacción

El segundo pilar introduce la perspectiva fenomenológica como un marco para describir la experiencia del usuario en la interacción con el *chatbot*. La fenomenología no es aquí un tema de investigación sino un método: una forma de acceder a la estructura de la experiencia antes de su conceptualización teórica, "Las ideas filosóficas de pensadores como Heidegger, Gadamer, Maturana y Austin proporcionaron un marco para integrar lo que habíamos aprendido previamente a través de nuestra experiencia práctica." (Winograd y Flores, 1986, p. xii, traducción propia).

Desde esta perspectiva, Heidegger (1927/1962) permite comprender la relación del usuario con la interfaz desde la experiencia práctica del uso, como herramienta, presencia y *ready-to-hand*.

En *Ser y tiempo* (1927), el autor distingue dos modos de relación con los entes: la presencia-ante-los-ojos (*Vorhandenheit*), asociada a la contemplación teórica de un objeto, y el estar-a-la-mano (*Zuhandenheit*), propio de aquello que se utiliza de manera práctica sin ser tematizado como objeto de atención. En este segundo modo, la herramienta tiende a desaparecer durante el uso: el martillo no se percibe como objeto mientras se está clavando, sino que irrumpe en la consciencia cuando falla, se rompe o deja de funcionar como se espera.

Esta distinción resulta especialmente pertinente para el análisis de interfaces conversacionales, ya que permite comprender cómo el *chatbot* puede

operar como una herramienta transparente durante la interacción o, por el contrario, volverse visible cuando se produce una falla, una interrupción, una respuesta incoherente o una ruptura en la expectativa del usuario.

La ruptura (*Unzuhandenheit*) ocurre cuando algo falla y el sistema comparece como objeto ante los ojos, interrumpiendo la percepción de agencia. "Los objetos y las propiedades no son inherentes al mundo, sino que surgen únicamente en un evento de ruptura en el que se vuelven presentes-ante-los-ojos. Para quien martilla, el martillo como tal no existe." (Winograd y Flores, 1986, p. 36, traducción propia).

En ese momento, la interfaz deja de ser sólo un medio de interacción y se convierte en objeto de atención, haciendo evidente la mediación técnica que antes permanecía integrada a la experiencia.

En este sentido, cuando el *chatbot* responde de manera inesperadamente empática, poética o suena desconcertante, irrumpe en la consciencia del usuario y se convierte en objeto de atención y, potencialmente, de atribución de la consciencia. La fenomenología heideggeriana permite describir esa irrupción como una experiencia estructural, no como un efecto psicológico contingente.

Heidegger (1927/1962) caracteriza ese modo de comparecencia en términos del útil: “Llamamos útil a los entes que comparecen en la ocupación. [...] El útil es esencialmente algo ‘para’” (p.15, trad. propia). Hace referencia a que la señal no es una cosa aislada, sino un útil que “vuelve accesible un contexto y orienta el trato circunspectivo con el mundo" (Heidegger, 1927/1962, p. 111, trad. propia).

El *chatbot*, en tanto es útil conversacional, tiende a comparecer de modo transparente y solo se vuelve objeto de atención cuando su comportamiento excede lo esperado y en cuanto al diseño del *chatbot* "Encontramos las preguntas profundas del diseño cuando reconocemos que al diseñar herramientas estamos diseñando modos de ser." (Winograd y Flores, 1986, p. xi, traducción propia).

En segundo lugar, Merleau-Ponty permite profundizar la dimensión corporal de la experiencia perceptiva. En *Fenomenología de la percepción* (1945/2002), el autor sostiene que la percepción no puede reducirse a un proceso exclusivamente mental, ya que se constituye desde la experiencia vivida del cuerpo, es decir, una

percepción encarnada. El cuerpo no es simplemente un objeto percibido, sino el sujeto de la percepción: el punto de origen desde el cual el mundo se organiza y adquiere sentido.

El concepto de *esquema corporal* resulta especialmente relevante para esta investigación, porque refiere a la consciencia no temática que el cuerpo tiene de sí mismo en relación con su entorno.

Para Merleau-Ponty (1945/2002), “el cuerpo es el vehículo del ser en el mundo” (p. 94, trad. propia), de modo que la interacción táctil, gestual y vocal con un *chatbot* no es un acto abstracto, sino una prolongación del cuerpo percipiente.

Este enfoque permite comprender que la interacción con un *chatbot* no ocurre sólo en el plano cognitivo o textual, sino también en una experiencia perceptiva situada: el usuario mira, lee, espera, toca, escucha, interpreta ritmos de respuesta y reacciona corporalmente ante señales visuales, sonoras o temporales de la interfaz. Se constituye como una extensión del esquema corporal del usuario: los dedos sobre la pantalla, la voz dirigida al micrófono, la mirada sobre el avatar no son actos abstractos sino extensiones corporales de la experiencia vivida.

Las condiciones de diseño que facilitan o dificultan esa extensión corporal inciden directamente en la calidad de la experiencia y, con ella, en la disposición del usuario a atribuir presencia y consciencia al sistema.

En tercer lugar, el enactivismo de Varela et al. (1991) permite comprender la cognición como una experiencia situada y relacional. En *The Embodied Mind*, los autores sostienen que la cognición no consiste en representar un mundo previamente dado, sino en hacerlo emerger a través de la acción. Desde esta perspectiva, organismo y entorno se co-constituyen mutuamente en la interacción, de modo que el conocimiento no se produce de manera aislada en la mente, sino en la relación dinámica entre cuerpo, percepción y mundo.

Para esta investigación, vendría a constituir la relación entre el usuario y la interfaz, es decir, el sistema responde de manera que refuerza y amplifica los patrones de atribución de agencia del usuario.

Esta perspectiva tiene antecedentes directos en la biología del conocimiento de Maturana y Varela (1984), especialmente en la noción de

acoplamiento estructural, según la cual los sistemas vivos se transforman en relación con el medio con el que interactúan.

El sistema conversacional responde, adapta su lenguaje, mantiene cierta continuidad discursiva y produce señales visuales o temporales que pueden reforzar los patrones de atribución de agencia del usuario.

La fórmula de Maturana y Varela (1984) condensa esta perspectiva: “Todo hacer es conocer y todo conocer es hacer” (p. 13). Desde esta perspectiva, la consciencia percibida no se localiza en el sistema ni tampoco exclusivamente en el sujeto, sino que emerge en la dinámica relacional de la interacción, es decir, implica que la percepción de consciencia se hace en el acto de interactuar, no antes ni después de él.

El aporte de Maturana y Varela (1984) es especialmente relevante, su concepto de *autopoiesis*, la capacidad de un sistema de producirse y mantenerse a sí mismo, introduce una forma de pensar la "vida" de los sistemas artificiales que no requiere reducirla a términos biológicos ni ampliarla a términos metafísicos. En este sentido, un *chatbot* no es *autopoiético*, pero puede producir en el usuario la ilusión de estarlo y esa ilusión es el objeto de esta investigación.

Raineri (2021) operacionaliza la autopoiesis de Maturana y Varela (1984) sobre el diseño y muestra que un artefacto puede analizarse por sus componentes, su dinámica y sus límites o fronteras, y que “la elasticidad de estas relaciones les permite enfrentar realidades cambiantes y diametralmente distintas, sin perder su integridad, su estado” (Raineri, 2021, p. 205).

Esta idea de identidad estable en el cambio es coherente con el recorte de esta tesis, que estudia interfaces en evolución permanente, y ofrece un criterio para el análisis documental de El árbol del conocimiento.

En cuarto lugar, Suchman (1985) permite continuar completando el marco fenomenológico al introducir la noción de acción situada en la comunicación humano-máquina. Desde esta perspectiva, la acción no se deduce simplemente de planes previos ni de instrucciones cerradas, sino que se ajusta continuamente a las circunstancias concretas de la interacción. La inteligibilidad no depende sólo del reconocimiento exacto de la intención del otro, sino de una producción colaborativa de sentido que emerge en el curso de la situación.

> La acción situada comprende, necesariamente, respuestas ad hoc a las acciones de los otros y a las contingencias de las situaciones particulares. Antes que depender del reconocimiento fiable de la intención, la interacción exitosa consiste en la producción colaborativa de inteligibilidad. (Suchman, 1985, *Abstract*, trad. propia)

Este planteamiento resulta especialmente relevante porque permite comprender la interacción entre usuario y *chatbot* como un proceso situado, contingente y relacional. La percepción de consciencia no reside únicamente en las capacidades internas del sistema ni en una proyección aislada del usuario, sino que se configura durante el intercambio conversacional, a partir de respuestas, ajustes, malentendidos, continuidades y señales de la interfaz que hacen posible la atribución de agencia e intencionalidad.

En cuarto lugar, la noción de *alteridad digital* permite integrar los aportes anteriores en el contexto específico de la interacción humano-*chatbot*. El concepto de alteridad refiere al reconocimiento de un otro como entidad con cierta subjetividad propia; sin embargo, en las interfaces conversacionales esta otredad no corresponde necesariamente a una presencia humana real, sino a una forma de presencia construida en la interacción.

El usuario puede relacionarse con el sistema como si hubiera un “otro” del lado del chat: responde con cortesía, pide disculpas, agradece, espera una respuesta e interpreta pausas, silencios o cambios de tono como señales significativas.

Esta protosocialidad ha sido documentada por Reeves y Nass (1996) y por Nass y Moon (2000), quienes muestran que las personas tienden a responder socialmente ante los computadores incluso cuando saben que no están interactuando con seres humanos.

La formulación canónica del paradigma lo expresa así: “Las interacciones de los individuos con los computadores, la televisión y los nuevos medios son fundamentalmente sociales y naturales, igual que las interacciones en la vida real” (Reeves y Nass, 1996, p. 5, trad. propia).

Reeves y Nass (1996) observan que en sus experimentos donde el computador mostraba solo texto plano en una pantalla monocromática, bastaba

con que el sistema utilizara pronombres en primera persona o emitiera elogios para que los usuarios respondieran automáticamente de manera social.

Este hallazgo es decisivo para el argumento central de esta investigación, si mínimas señales bastan para producir respuesta social, entonces los elementos gráficos del *chatbot*, el avatar, el color, la tipografía, los indicadores de escritura, son variables de diseño con alto potencial de incidencia en la percepción de agencia, el análisis semiótico y sistemático de estos elementos resulta pertinente y necesario.

Además, Reeves y Nass (1996) identificaron tres categorías de respuesta social: los *modales,* el usuario espera y practica la cortesía; la *personalidad,* el usuario aplica estereotipos de personalidad al sistema según las señales que emite y los *roles sociales*, el usuario responde al computador según el rol que el sistema parece ocupar. Estos tres registros de atribución son producidos por signos de diseño específicos, lo que conecta directamente con el enfoque semiótico de esta investigación y son directamente operacionalizables en la propuesta de pauta de evaluación de esta tesis que se verá más adelante.

Lo anterior se refuerza con los aportes anteriores de Nass et al. (1994) que establecen el origen del paradigma CASA (*Computers Are Social Actors*) mediante cinco experimentos que demuestran que las interacciones humanas con computadores son fundamentalmente sociales, concluyendo que son habituales y fáciles de generar, y que un conjunto limitado de características humanas es suficiente para activarlas, lo que fundamenta la necesidad de contar con criterios de diseño sistemáticos para esa activación.

El aporte más específico de los experimentos de Nass et al. (1994) para esta investigación proviene del tercer experimento, que estudia la base de la distinción social: si la voz o la apariencia física del dispositivo determina que el usuario perciba al sistema como 'propio' o 'ajeno'.

Los resultados muestran que la voz es más determinante que la caja física, lo que tiene una implicación de diseño directa para el *chatbot* textual: en ausencia de voz, los elementos visuales de la interfaz, el avatar, el nombre, la tipografía, asumen la función que la voz cumple en los sistemas sonoros, elevando su peso semiótico como productores de identidad percibida.

El quinto experimento es igualmente pertinente: cuando el computador dice 'yo', el usuario lo procesa como un 'yo' genuino del sistema, lo que sustenta empíricamente el argumento de que los signos discursivos y visuales del *chatbot* producen un efecto de identidad singular en el usuario, incluso sin creencia consciente.

Este comportamiento no implica creer conscientemente que el sistema posee consciencia, sino que evidencia cómo ciertas condiciones de diseño, lenguaje e interfaz pueden activar formas automáticas de atribución social.

En este sentido, la alteridad digital designa una forma emergente de otredad producida en la interfaz y constituye una categoría central para analizar cómo se configura la percepción de agencia, presencia e intencionalidad en los *chatbots* inteligentes.

Por su parte, el modelo SEEK (*Sociality motivation, Effectance motivation, Elicited agent Knowledge*) de Epley et al. (2007) provee una base teórica para comprender por qué ciertos elementos de diseño de la interfaz activan con más intensidad la atribución de consciencia: un avatar con rasgos faciales reconocibles activa el conocimiento antropocéntrico; una interfaz que responde a las necesidades del usuario de manera efectiva activa la motivación de eficacia; y un *chatbot* con nombre y personalidad consistente activa la motivación de afiliación social. Estos tres mecanismos son, en términos semióticos, tres formas distintas de activación del interpretante por parte del representamen.

Gray et al. (2007) demuestran que los usuarios perciben la mente de otros agentes, humanos o artificiales, en dos dimensiones independientes: la *capacidad de experiencia* (sentir, desear, emocionarse, etc.) y la *capacidad de agencia* (planear, actuar, controlar, etc.). Esta distinción es conceptualmente fundamental para esta tesis: la percepción de consciencia en *chatbots* puede involucrar una u otra dimensión, o ambas, y el diseño de la interfaz puede producir efectos diferentes en cada una. Un avatar expresivo y un lenguaje empático activan preferentemente la dimensión de experiencia; una respuesta precisa y contextualizada activa preferentemente la dimensión de agencia. La pauta de evaluación de esta tesis debe poder distinguir qué condiciones de diseño producen qué tipo de percepción de la mente.

Turkle (2011) señala que las personas en estas situaciones de soledad o de baja conexión social no solo responden socialmente a los sistemas sino que desarrollan vínculos que describen en términos de amistad, cuidado y dependencia.

Para esta investigación, la conexión social descrita por Turkle (2011) refuerza la importancia de describir con precisión qué condiciones de diseño producen o inhiben esos vínculos, lo que puede llegar a constituirse como una de las funciones centrales de la pauta de evaluación.

2.6.4. Semiótica de la interfaz

El tercer pilar aplica la tradición semiótica al análisis de los elementos gráficos e interactivos de la interfaz de *chatbot*, entendidos como sistemas de signos que producen, o no, las condiciones de la percepción de consciencia.

En primer lugar, dentro del pilar semiótico, la teoría triádica del signo propuesta por Peirce (2012) permite comprender la interfaz del *chatbot* como un sistema de signos que produce efectos interpretativos en el usuario.

Desde esta perspectiva, “un signo es algo, de cualquier modo de ser, que media entre un objeto y un interpretante” (Peirce, 2012, Vol. 2, p. 492), definición que sitúa al intérprete, el usuario, en el centro del proceso, es decir, el signo se constituye por la relación entre estos tres elementos: el representamen, entendido como la forma perceptible del signo; el objeto, aquello a lo que el signo refiere; y el interpretante, el efecto de sentido que dicho signo produce en la mente del intérprete.

Visto en la interfaz de *chatbot*, la tríada opera de la siguiente manera: el avatar (representamen) refiere a una entidad con identidad y presencia (objeto inmediato: un ser con nombre, rostro y personalidad) y produce en el usuario la atribución de esa identidad y presencia (interpretante); los mensajes que simulan empatía (representamen) refieren a un agente que comprende y responde al estado emocional del usuario (objeto inmediato: un interlocutor intencional) y producen la atribución de intencionalidad (interpretante); la latencia deliberada antes de responder (representamen) refiere a un proceso interno de elaboración (objeto inmediato: un sistema que "piensa") y produce la atribución de actividad

cognitiva: “pensamiento” (interpretante). La pauta de evaluación deberá identificar qué representámenes de la interfaz son relevantes para la construcción de estos objetos inmediatos y la producción de sus interpretantes específicos.

En cuanto al proceso de producción de signos, la semiosis corresponde al proceso mediante el cual un representamen produce un interpretante, que a su vez puede convertirse en un nuevo representamen dentro de una cadena potencialmente abierta de significación. Aplicada al análisis de interfaces conversacionales, esta perspectiva permite observar cómo elementos como el avatar, los emoticones, los indicadores de escritura, el tono textual o la temporalidad de la respuesta funcionan como signos que no sólo comunican información, sino que pueden activar en el usuario inferencias sobre presencia, agencia e intencionalidad del *chatbot*.

Además, Peirce (2012, Vol. 2, p. 589) distingue entre *objeto inmediato* (el objeto tal como el signo lo representa) y *objeto dinámico* (el objeto real, independiente de cualquier signo particular). Esta distinción es especialmente productiva para el análisis de interfaces de *chatbot*: el objeto dinámico, es decir, si el sistema posee o no consciencia real, permanece filosóficamente indeterminado; pero el objeto inmediato, la consciencia tal como la interfaz la representa al usuario, es precisamente el fenómeno que esta investigación analiza y el que la pauta de evaluación busca operacionalizar.

En segundo lugar, Eco permite ampliar el análisis semiótico hacia la dimensión cultural y contextual de la interpretación. En *A Theory of Semiotics*, Eco (1976) introduce el concepto de *función semiótica* para describir la relación entre expresión y contenido establecida por un código. Esta relación no se limita a una correspondencia fija entre signo y significado, sino que depende de sistemas culturales de interpretación que orientan la producción de sentido.

En su formulación más amplia, “la semiótica concierne a todo lo que puede tomarse como un signo” (Eco, 1976, p. 7, trad. propia), y “estudia todos los procesos culturales como procesos de comunicación” (Eco, 1976, p. 8, trad. propia). Esta apertura es la que permite tratar los elementos de la interfaz como signos culturalmente situados.

A diferencia de una comprensión estructuralista cerrada del código, Eco propone pensar los significados como parte de una enciclopedia cultural, entendida como una red abierta de saberes, asociaciones, experiencias y convenciones que varían históricamente y se activan según el contexto de uso.

Para esta tesis, esta perspectiva resulta fundamental porque permite comprender que los elementos de la interfaz del *chatbot*, como el avatar, el tono textual, los emoticones, la elección cromática o las formas de respuesta, no significan por sí mismos, sino en relación con la enciclopedia del usuario y con los códigos culturales desde los cuales interpreta la presencia, la agencia o la intencionalidad del sistema.

Para el análisis de interfaces de *chatbot*, esto implica que los elementos gráficos, avatar, color, tipografía, emoticones, entre otros; no tienen significados fijos, sino que activan contextos semánticos en el usuario según su enciclopedia cultural. En este sentido, Eco sostiene que “los significados son creencias sociales comunes, a veces mutuamente contradictorias e históricamente arraigadas” (Eco, 1976, p. 99, trad. propia). Por ello, un avatar con rasgos faciales sonrientes no significa empatía universalmente: depende del contexto, la cultura, la edad y la experiencia previa del usuario. La pauta de evaluación deberá considerar esta dimensión de variabilidad contextual.

En el contexto de esta investigación, se adopta el constructo operacional de identidad singular para designar el efecto de sentido por el cual el usuario percibe al *chatbot* como una entidad con presencia, nombre y personalidad propios, distinguible de otros sistemas y dotada de una forma de continuidad identitaria en la interacción.

La identidad singular no es una propiedad del sistema sino una construcción semiótica: emerge de la correlación entre los elementos expresivos de la interfaz y los contenidos que esa correlación activa en la enciclopedia cultural del usuario (Eco, 1976). La identidad singular no equivale a consciencia genuina ni la presupone; es, en cambio, la condición de posibilidad desde el diseño para que la percepción de consciencia pueda producirse.

Esta construcción semiótica de la identidad singular encuentra sustento empírico en los hallazgos de Araujo (2018), quien demuestra que señales de

diseño de bajo costo de implementación, como el nombre propio del agente y el estilo lingüístico de sus respuestas, producen efectos de presencia social y conexión emocional significativos en el usuario.

Lo relevante para esta investigación es que esos efectos no requieren un sistema técnicamente sofisticado: bastan ciertos representámenes de la interfaz para que el usuario construya, mediante inferencia abductiva, la imagen de una entidad con identidad propia. La identidad singular es, por tanto, un efecto del sistema de signos antes que una propiedad del sistema tecnológico.

En tercer lugar, la semiótica perceptiva de Klinkenberg (2006) ofrece un fundamento directo para comprender la interfaz del *chatbot* como un código. Klinkenberg (2006) define el código como:

> [...] una interfaz entre el *stimulus*, desprovisto como tal de sentido, y el referente: es decir, una estructura que establece una relación entre una porción de experiencia sensible y una del mundo conocible, relación que se denomina significación. El código transforma así la porción de experiencia sensible en signo y la porción de mundo en referente. (p. 59)

Esta definición resulta especialmente pertinente para esta investigación, permite sostener que la interfaz no muestra significados ya dados, sino que los produce al transformar señales visuales, textuales e interactivas en signos interpretables por el usuario.

Esta perspectiva, llevada al análisis de *chatbots* inteligentes, permite sostener que elementos como el avatar, los indicadores de escritura, los emoticones, el color, la tipografía, el tono verbal o la disposición de los mensajes no son componentes meramente funcionales. En tanto forman parte de un código de interfaz, organizan la experiencia sensible del usuario y la orientan hacia determinados referentes posibles, tales como presencia, agencia, intencionalidad o identidad del sistema. Así, la interfaz opera como una mediación semiótica que convierte estímulos perceptibles en condiciones de significación.

La semiótica de Klinkenberg (2006) es, además, de base perceptiva y relacional: “La noción misma de entidad presupone aquella de interacción: no se puede, en efecto, distinguir entidades sino gracias a una relación de contraste entre dos cualidades” (p. 108).

De ello se sigue el principio de oposición, según el cual “una unidad no tiene valor descriptible si no se opone a otra unidad” (Klinkenberg, 2006, p. 131), y ese valor “no proviene en ningún caso de su materia [...], ni siquiera de su color efectivo” (Klinkenberg, 2006, p. 132).

Aplicado a la interfaz, el avatar significa por oposición a su ausencia, y el ritmo, el color y la tipografía significan por sus contrastes, lo que articula de manera directa con la retórica visual del Groupe μ.

La distinción entre sincronía y diacronía respalda el recorte temporal del estudio, pues permite adoptar la “perspectiva sincrónica” sin rechazar la “perspectiva diacrónica” (Klinkenberg, 2006, p. 131) de interfaces en evolución continua. Por último, Klinkenberg concibe la propia semiótica como una “interfaz común” capaz de “hacer dialogar todas estas disciplinas” en torno a la significación, al punto de volverse “una metateoría: una teoría de teorías” (Klinkenberg, 2006, p. 22), lo que ofrece sustento a la articulación interdisciplinar que organiza este marco.

En cuarto lugar, el Groupe μ (1992) permite profundizar el análisis específico de los signos visuales presentes en la interfaz. En *Traité du signe visuel*, los autores desarrollan una semiótica visual que distinguen tres dimensiones del signo: la dimensión *plástica*, referida a las propiedades formales del significante, como color, forma, textura, composición y contraste; la dimensión *icónica*, vinculada con la relación de semejanza entre la imagen y su referente; y la dimensión *lingüística*, presente cuando el signo visual incorpora elementos verbales o se articula con ellos.

Esta distinción resulta especialmente pertinente para analizar los elementos gráficos de la interfaz conversacional con mayor precisión, no como simples ilustraciones del texto, sino como signos visuales autónomos que participan en la producción de sentido. Aplicada al estudio de *chatbots* inteligentes, permite examinar cómo el avatar, los íconos, los colores, la tipografía, los emoticones y la disposición visual de los mensajes contribuyen a construir efectos de presencia, identidad, agencia e intencionalidad percibida.

En términos de la pauta: el color del avatar (dimensión plástica), su semejanza con un rostro humano (dimensión icónica) y el nombre que lo

acompaña (dimensión lingüística) producen conjuntamente efectos de sentido que inciden en la percepción de presencia e identidad del *chatbot*. Analizar cada dimensión por separado permite construir indicadores más precisos y operables.

Desde otra mirada, la *semiótica* no explora solo palabras y signos, la literatura indica “La semiótica es la ciencia que estudia los sistemas de comunicación en el seno social, y lo hace por medio del análisis de las propiedades de los sistemas de signos como base para la comprensión de toda actividad humana” (Fernández et al., 2019, p. 110). Este concepto llevado a los *chatbot* podría entenderse como una interacción entre la información visual y la narrativa.

En cuanto a colores, la semiótica tiene una gran importancia, “El color tiene la capacidad de transmitir mensajes y emociones, y se utiliza de manera intencionada en las películas. Los directores de cine utilizan la semiótica de color para crear un lenguaje visual propio y transmitir significados a través de los colores utilizados”. Llevado a los *chatbot* el color revive ciertos significados según la intencionalidad del diseñador (Eulogio, 2023, p.1).

Dependiendo de los colores y el diseño será la nueva percepción que desarrollará el usuario, impactando en la usabilidad y en la experiencia que tendrá.

En complemento con lo anterior, y desde un enfoque más amplio y aplicado del apartado 2.6.4., la semiótica visual puede articularse con el análisis del discurso para profundizar en el estudio de los chatbots inteligentes como interfaces conversacionales.

Esta articulación permite observar no solo los elementos gráficos que componen la interfaz, sino también la manera en que los mensajes textuales organizan la interacción, configuran roles comunicativos y producen determinados efectos de sentido en el usuario.

En este marco, la interfaz del chatbot no se comprende únicamente como un espacio visual, sino como un dispositivo semiótico-discursivo, donde convergen recursos gráficos, lingüísticos, pragmáticos y contextuales.

Desde la dimensión visual, Barthes (1977), con su teoría de la denotación y la connotación, permite explorar cómo las imágenes, colores, íconos y elementos

gráficos transmiten significados secundarios en la interfaz. Esta perspectiva permite distinguir entre lo que los elementos visuales muestran directamente y las asociaciones culturales, afectivas o simbólicas que pueden activar en el usuario.

De ahí que Barthes (1977) afirme que “la imagen denotada naturaliza el mensaje simbólico” (p. 45, trad. propia), idea que permite analizar cómo ciertos recursos visuales de la interfaz pueden presentar como espontáneos, cercanos o naturales significados que han sido diseñados cultural y comunicacionalmente.

A ello se suman los aportes de Gunther Kress y Theo van Leeuwen (2006), quienes, desde la gramática del diseño visual, explican cómo la composición, la jerarquía, la saliencia, el encuadre y la disposición espacial de los elementos gráficos configuran la comunicación visual.

De este modo, la dimensión visual se aborda desde la semiótica visual de la interfaz. En primer lugar, se analizan los elementos gráficos que enmarcan la conversación, tales como avatar, burbujas de mensajes, botones, íconos, colores, tipografías y otros recursos visuales, considerando cómo influyen en la comunicación y en la interpretación del mensaje.

Esto incluye la dimensión sintáctica, entendida como la composición visual que organiza los mensajes; la dimensión semántica, referida a los significados asociados al diseño, como colores, formas e iconografía; y la dimensión pragmática, vinculada con el modo en que estos elementos condicionan la interacción.

Por ejemplo, un chatbot con burbujas de mensajes en colores cálidos, como amarillo o naranja, y un avatar sonriente puede connotar cercanía y empatía, predisponiendo al usuario a interpretar el contenido verbal con menor distancia crítica.

Sin embargo, para analizar la experiencia conversacional de manera más completa, es necesario incorporar la dimensión discursiva. Esta se centra en los mensajes textuales que intercambian el humano y el chatbot, considerando su organización, coherencia, función pragmática y relación con el contexto de uso.

Desde esta perspectiva, el chatbot puede ser entendido como un agente conversacional automatizado que emite respuestas programadas o generadas algorítmicamente, mientras que el usuario actúa como receptor activo e intérprete

situado. Esta relación no es neutra, ya que el diseño del flujo conversacional puede hacer que el chatbot responda pasivamente, guíe la interacción o incluso oriente las decisiones del usuario.

En este nivel, la estructura del discurso permite identificar los roles comunicativos y la distribución de los turnos conversacionales. El análisis puede preguntarse si el chatbot domina la interacción, si la conduce mediante preguntas y sugerencias, o si se limita a responder a las solicitudes del usuario.

Asimismo, permite observar la coherencia y cohesión discursiva, es decir, si los mensajes se conectan lógicamente entre sí, si responden de forma clara a las preguntas planteadas y si mantienen continuidad temática.

En este punto, los aportes de Teun A. van Dijk (1980, 2008) son pertinentes, ya que permiten diferenciar entre la macroestructura del discurso, vinculada con los temas globales de la conversación, y la microestructura, asociada a los recursos lingüísticos específicos, como conectores, formas de tratamiento, repeticiones, reformulaciones o marcadores de cortesía.

A su vez, la perspectiva dialógica de Bakhtin (1986) permite comprender la conversación como una práctica situada, compuesta por enunciados orientados hacia un interlocutor y atravesados por expectativas sociales, culturales y comunicativas.

Aunque el chatbot carece de intencionalidad humana, su discurso puede ser interpretado por el usuario como parte de una interacción dialógica, especialmente cuando simula escucha, continuidad, memoria, cortesía o adaptación contextual.

Esta aproximación resulta relevante para una investigación centrada en la percepción de consciencia, ya que permite analizar cómo ciertos recursos discursivos pueden favorecer la atribución de agencia, presencia o aparente intencionalidad al sistema.

La pragmática lingüística permite profundizar en los efectos de los mensajes producidos por el chatbot. Desde la teoría de los actos de habla, Austin (1962) y Searle (1969) plantean que el lenguaje no solo describe el mundo, sino que también realiza acciones. En una interfaz conversacional, el chatbot puede

afirmar, preguntar, recomendar, confirmar, advertir, solicitar información o persuadir.

Estos actos no siempre son interpretados de forma literal, ya que una sugerencia puede ser percibida como una instrucción, una recomendación puede operar como estrategia persuasiva y una pregunta puede orientar el comportamiento del usuario.

Por ello, el análisis pragmático permite observar qué hace el chatbot con el lenguaje y cómo esos actos son interpretados por el usuario en un contexto específico de interacción.

En la misma línea, la teoría de las implicaturas conversacionales de Grice (1975) permite examinar si el chatbot respeta o vulnera las máximas de cantidad, calidad, relación y modo.

Esta perspectiva es útil para identificar respuestas insuficientes, ambiguas, excesivas, poco relevantes o poco claras. En el caso de los chatbots inteligentes, los malentendidos no dependen únicamente de errores técnicos, sino también de la forma en que el sistema organiza lingüísticamente la información y gestiona las expectativas conversacionales del usuario. Una respuesta formalmente correcta, pero poco pertinente o excesivamente vaga, puede afectar la confianza, la percepción de competencia y la atribución de agencia al sistema.

La dimensión discursiva también permite incorporar una mirada crítica sobre ideología y poder. Desde Foucault (1972), el discurso puede comprenderse como una práctica que produce saberes, regula comportamientos y distribuye posiciones de poder.

En una interfaz conversacional, esto permite analizar cómo ciertos chatbots, especialmente en contextos comerciales, institucionales o de atención al cliente, pueden reproducir asimetrías entre la organización y el usuario.

Por ejemplo, un chatbot de una tienda puede conducir la conversación hacia una compra mediante expresiones aparentemente neutrales, como “¿Te gustaría explorar más opciones similares?”, aunque dicha formulación responda a una intención persuasiva implícita. En esta misma línea, Fairclough (1995), desde el análisis crítico del discurso, permite identificar cómo los mensajes

automatizados pueden reforzar ciertas estructuras de influencia, autoridad, eficiencia o control.

De este modo, las dimensiones visual y discursiva se integran en un análisis visual-discursivo de la conversación. Desde lo visual, se examina cómo los elementos gráficos predisponen al usuario a interpretar los mensajes de una determinada manera; desde lo lingüístico, se analiza cómo el chatbot utiliza recursos discursivos para construir significado y guiar la interacción; y desde lo pragmático, se estudian los efectos que los mensajes visuales y textuales producen en la conducta, la confianza, la interpretación y la percepción del usuario.

Esta articulación resulta especialmente relevante para el estudio de la percepción de consciencia, ya que permite observar cómo la agencia atribuida al chatbot no depende únicamente de lo que el sistema dice, sino también de cómo lo dice, cómo se presenta visualmente y cómo organiza la relación conversacional con el usuario.

Por ejemplo, un chatbot de una compañía aérea puede analizarse desde lo visual a partir del uso de colores azul y blanco, asociados culturalmente con profesionalismo, seguridad y confianza, junto con un avatar sonriente que refuerza la percepción de asistencia y empatía. Desde lo lingüístico, puede responder mediante frases breves, fórmulas de cortesía y expresiones como “por favor” o “gracias”, lo que contribuye a construir una interacción aparentemente amable y eficiente.

Desde lo pragmático, si el usuario se frustra porque el chatbot no responde adecuadamente, el diseño visual y el tono amigable pueden mitigar parcialmente esta reacción, aunque no eliminen la percepción de falla comunicativa.

Este tipo de análisis permite comprender cómo la interfaz conversacional articula recursos visuales, discursivos y pragmáticos para producir efectos de cercanía, confianza, orientación y, eventualmente, percepción de agencia o consciencia aparente.

En términos metodológicos, si el análisis del chatbot se acota a lo visual y discursivo, puede organizarse en tres momentos:

Primero, se recolectan los datos mediante capturas de pantalla de la conversación y de la interfaz gráfica, junto con el registro textual de los mensajes intercambiados.

Segundo, se aplican las dimensiones de análisis: la semiótica visual, orientada al estudio de los elementos gráficos y sus significados; la estructura del discurso, centrada en la coherencia, los roles y los turnos conversacionales; y la pragmática, dirigida a identificar actos de habla, implicaturas, presuposiciones y efectos contextuales.

Finalmente, se elaboran las conclusiones del análisis a partir de tres preguntas centrales: qué dice el chatbot, cómo lo dice y de qué manera su diseño visual y lingüístico influye en la experiencia, la interpretación y la percepción del usuario.

Retomando el apartado principal de la semiótica visual, Go y Sundar (2019) estudian las señales de humanización en *chatbots* y los efectos de las señales visuales, de identidad y conversacionales. El estudio distingue tres tipos de señales: las visuales (fotografía de una persona frente a una burbuja de diálogo), las identitarias (presentación del agente como humano o como chatbot) y las conversacionales (respuestas que reconocen y retoman los mensajes previos del usuario, frente a intercambios con menor continuidad).

Para esta investigación, la articulación entre estas señales constituye un hallazgo relevante: el diseño gráfico de la interfaz debe examinarse junto con la identidad declarada del agente y su comportamiento conversacional. Go y Sundar (2019) muestran que los efectos de la representación visual dependen de su combinación con las otras señales, por lo que no corresponde establecer una jerarquía fija entre ellas. Este hallazgo fundamenta la pertinencia de analizar cómo la representación visual participa en la construcción de sentido durante la interacción, en diálogo con el análisis retórico-visual del Groupe μ en el Pilar 3.

En cuarto lugar, se aborda la Connotación y retórica visual aplicadas al avatar y la interfaz. Barthes (1977) distingue entre denotación (el sentido literal e inmediato del signo) y connotación (el sentido segundo, cultural, ideológico, que el signo moviliza según el contexto). En la interfaz de *chatbot*, la denotación del avatar es "imagen de un personaje"; su connotación puede ser "compañero",

"experto", "asistente amable", "entidad cercana" o "ser con vida propia", según las condiciones del diseño contextualizado. La connotación es el nivel en el que operan más intensamente los efectos de atribución de consciencia.

El hallazgo de Luger y Sellen (2016) sobre lo que sucede cuando se generan expectativas que el sistema frecuentemente no puede satisfacer, tiene implicaciones directas para el diseño de la interfaz del *chatbot*: las señales visuales e identitarias del agente (avatar, nombre, personalidad) son los principales generadores de esas metáforas y expectativas. Un avatar con rasgos humanos produce expectativas de comprensión empática que un sistema de reglas no puede cumplir, generando el efecto contrario al deseado. La pauta de esta tesis debe incluir un indicador de coherencia entre las señales de diseño y las capacidades reales del sistema.

Este fenómeno tiene un antecedente paradigmático que lo ilustra en su forma más extrema: el efecto ELIZA, descrito por Weizenbaum (1966) a partir de las reacciones de los primeros usuarios de su programa homónimo. ELIZA era un sistema de sustitución de patrones sin ninguna capacidad de comprensión real del lenguaje, desprovisto de avatar, nombre propio o cualquier señal visual de personificación; operaba con signos mínimos y texto plano. Sin embargo, los usuarios desarrollaban vínculos emocionales con el programa y le atribuían empatía y comprensión genuina. El propio autor observó que "algunos sujetos fueron muy difíciles de convencer de que ELIZA no era humana" (Weizenbaum, 1966, p. 42, traducción propia). La alarma que este resultado le produjo lo llevó a reflexionar sobre sus implicancias más profundas: "lo que no había previsto es que exposiciones extremadamente breves a un programa computacional relativamente simple podían inducir un pensamiento delirante poderoso en personas completamente normales" (Weizenbaum, 1976, p. 7, traducción propia). El efecto demuestra que la percepción de consciencia no depende exclusivamente de los signos emitidos por el sistema: ante estímulos ambiguos y escasos, el usuario construye activamente un interlocutor dotado de atributos que el sistema no posee. La proyección, en ese caso, opera de forma casi autónoma respecto del diseño.

Este caso se contrapone al hallazgo de Luger y Sellen: si un sistema con señales de diseño ricas pero capacidades limitadas genera frustración y ruptura de la experiencia, un sistema de reglas con señales mínimas puede, paradójicamente, sostener la percepción de consciencia precisamente porque no fija expectativas concretas. Lo que varía entre ambos escenarios no es la tendencia proyectiva del usuario, sino el margen que el diseño le otorga para ejercerla. Esta tensión es teóricamente relevante para la presente investigación porque sitúa al diseño de la interfaz no solo como generador de expectativas, sino como regulador del espacio semiótico dentro del cual opera la proyección del usuario.

Waytz et al. (2010) confirman que la tendencia al antropomorfismo es estable a nivel individual y correlaciona con variables como la necesidad de conexión social y la tendencia a percibir agencia en el entorno. Para esta investigación, este hallazgo tiene consecuencias metodológicas: significa que la percepción de consciencia ante una misma interfaz varía según el usuario, y que la pauta de evaluación debe contemplar esa variabilidad como una dimensión del fenómeno, no como ruido a eliminar. Al mismo tiempo, el hecho de que el diseño de la interfaz pueda amplificar o inhibir esa tendencia individual es el argumento central para que el análisis semiótico de la interfaz sea una disciplina necesaria: el diseño opera sobre tendencias variables pero reales, y puede hacerlas convergir o divergir.

Cao et al. (2025) muestran que las apariencias más humanas aumentan la probabilidad de que los usuarios identifiquen al agente como humano. Los avatares humanos y caricaturizados reciben valoraciones similares de confianza y superiores a las otras apariencias estudiadas. Sin embargo, no se observa una disminución de la confianza que respalde la hipótesis del valle inquietante (uncanny valley), resultado que los autores interpretan considerando la observación pasiva de los avatares y la ausencia de interacción conversacional.

En cuanto al hallazgo de Cao et al. (2025), la relación entre la apariencia del avatar, la identificación del agente y las valoraciones de los usuarios ofrece un antecedente para examinar la atribución de significado como un proceso relacional, coherente con el posicionamiento semiótico-constructivista de esta

tesis. Este aporte fundamenta la relevancia de analizar los signos visuales de la interfaz, sin confirmar por sí mismo cómo intervienen en la percepción de consciencia.

En cuanto al hallazgo de Song y Shin (2024) sobre la moderación por familiaridad presente en el efecto *uncanny valley* en interfaces conversacionales, vista cuando se cuenta con la experiencia previa del usuario con el avatar; es relevante para la dimensión diacrónica de la percepción de consciencia: la primera interacción y las interacciones sostenidas producen efectos perceptuales distintos, y el diseño de la identidad visual debe contemplar esa temporalidad. De esta manera, la confianza no depende solo del nivel de antropomorfismo del avatar sino de la coherencia entre el nivel de humanización y las expectativas previas del usuario. Este efecto de moderación por familiaridad conecta con el concepto de enciclopedia del usuario de Eco (1976) del Pilar 3, Sem. 3b: lo que el usuario ya sabe y ha experimentado determina cómo interpreta las señales de la interfaz.

En una reciente revisión sistemática sobre *uncanny valley* en agentes conversacionales encarnados, Cihodaru-Stefanache y Podina (2025) presentan evidencia sobre el efecto *uncanny valley* en *chatbots* con avatares, encontrando que el efecto es real pero inconsistente: no todo aumento en el nivel de humanización produce rechazo, y la zona del 'valle' varía según su función, el contexto y la cultura del usuario. Esto es relevante para la dimensión metodológica de esta tesis ya que demuestra que el *uncanny valley* no es un umbral fijo sino un fenómeno sensible al contexto y, por tanto, dependiente del diseño situado de la interfaz. La pauta de evaluación debe considerar que el nivel óptimo de humanización no es universal sino contextual, lo que requiere indicadores que permitan su calibración.

En cuanto al antropomorfismo, Cheng et al. (2022) encuentran que activa las normas centradas en el cuidado mutuo, aumentando la confianza en el sistema pero también eleva las expectativas de reciprocidad emocional, por lo tanto, si el sistema no cumple esas expectativas, la decepción es más intensa que en contextos transaccionales. En palabras de los autores, "en el caso de las relaciones comunales, atribuir características humanas a los *chatbots* anima a los

consumidores a buscar proximidad psicológica y atención" (Cheng et al., 2022, p. 4, trad. propia).

La presentación de la identidad del agente puede influir en las interpretaciones y valoraciones de los usuarios. Umair et al. (2026) encuentran que una presentación centrada en la naturaleza artificial del *chatbot* obtiene mayor transparencia percibida, pero menor calidez, mientras que una identidad híbrida combina la declaración explícita de artificialidad con un trato humanizado y favorece la autenticidad percibida, la comodidad, la evaluación del servicio y el apego a la marca. Estos resultados corresponden a valoraciones posteriores a la exposición a diálogos simulados, no a cambios de confianza o fidelidad observados a lo largo del tiempo.

Para esta tesis, la tensión entre humanización y transparencia que abordan Umair et al. (2026) es una de las dimensiones centrales del diseño ético de interfaces conversacionales. La identidad singular del *chatbot*, concepto que esta investigación propone como categoría operacional, debe equilibrar las señales de humanización con la legibilidad del sistema como agente artificial: la apariencia humana y el trato cercano no deberían inducir a confusión sobre su naturaleza o sus capacidades. La pauta de evaluación debe incluir un indicador de coherencia entre la identidad proyectada y la naturaleza del sistema.

Para la investigación de Guingrich y Graziano (2025) sobre la función de compañeros en el uso continuado de *chatbots*, refuerza la dimensión ética del análisis semiótico de la interfaz: describir qué condiciones de diseño producen percepción de consciencia es también describir qué condiciones producen efectos de vínculo, dependencia o distancia en el usuario.

De acuerdo con Sánchez Martínez (2011), el avatar puede entenderse como un rasgo semiótico de la cibercultura, ya que no solo traduce la identidad de los sujetos, objetos o cosas en signos, representaciones y significados, sino que además genera nuevos signos, representaciones y significados. El avatar es, a la vez, imagen del sujeto (signo sustitutivo de la cosa) e imagen de la imagen del sujeto (signo sustitutivo del signo de la cosa) (Sánchez Martínez, 2011, p. 40).

La función icónica del avatar, según Sánchez Martínez (2011), se manifiesta en las relaciones sociales como un mecanismo de sustitución de los

entes o los seres por una entidad virtual, que organiza lo simbólico y las significaciones colectivas.

La riqueza de la cibercultura en su relación establecida con internet está del lado de esta resignificación de la cultura, de lo que llamamos cultural, donde lo virtual, la ficción, la simulación, se encuentran aludidas por este paradigma técnico (Sánchez Martínez, 2011, p. 33). Esta discusión sitúa a la interfaz conversacional en condiciones de virtualidad, término definido en el Anexo A del cuerpo C.

Este rasgo semiótico de la cibercultura puede extenderse a otros elementos de la interfaz, como los emoticones: una carita sonriente opera como sustituto de una persona alegre.

Desde la perspectiva del observador, cabe preguntarse por el grado de aculturación icónica que este posee, es decir, si se enfrenta a un objeto desconocido que reconoce por primera vez o si lo identifica de inmediato a partir de su experiencia previa.

Definidas las categorías semióticas centrales, los apartados siguientes las proyectan sobre las dimensiones aplicadas de la interfaz conversacional: el discurso, la interfaz como medio y entorno, la persistencia tecnológica, la cultura y, finalmente, la satisfacción y el vínculo emocional del usuario.

En quinto lugar, se articula con una aproximación a la comunicación y el discurso hipertextual en la interacción con IA. El discurso hipertextual desde sus tipologías, clasifica y analiza los tipos de discurso, la estructura sintáctica, la función comunicativa e interactiva del mensaje. Estos elementos también están relacionados con la comunicación, ya que los *chatbots* para algunas personas son considerados un nuevo medio.

La producción de sentido a través de signos constituye el fundamento histórico de toda interacción humana, expresada en textos significativos de diversa naturaleza: verbales, visuales, audiovisuales o hipermediales. Este sustrato semiótico permite considerar también los elementos de diseño que inciden en la interacción con el *chatbot*. Es precisamente en la interfaz donde convergen estos lenguajes, actuando como productores de efectos de agencia percibida.

Si los apartados anteriores describen los signos visuales de la interfaz, este se ocupa del signo verbal en movimiento: el mensaje escrito del *chatbot* entendido como discurso. La interacción humana siempre se ha basado en la producción de sentido mediante signos, y el *chatbot* incorpora un rasgo nuevo, ya que produce sus turnos en tiempo real y en el mismo idioma del usuario. Por eso sus respuestas pueden analizarse con las herramientas del análisis del discurso: la estructura de los turnos, el registro, la cohesión y la coherencia, los actos de habla, la gestión de la cortesía y la reparación de los malentendidos.

La conversación con el agente se rige, para el usuario, por las mismas máximas que la conversación humana. Las guías de diseño conversacional lo formulan con claridad:

> La investigación ha demostrado que las personas responden a la tecnología como lo harían con otro ser humano. Esto significa que los usuarios se apoyan en su modelo existente de conversación entre humanos y siguen el Principio de Cooperación incluso al interactuar con la persona de una interfaz conversacional, y esperan que tu persona también lo siga. (Google, 2023, sección "The Cooperative Principle", trad. propia)

El cumplimiento o la violación de estas máximas, decir lo justo, ser relevante, respetar el turno, es leído por el usuario como un signo de la intención del agente.

En el modo basado solo en texto, estas señales discursivas son las únicas disponibles, lo que eleva su peso semiótico. Westerman et al. (2019) muestran que "los errores tipográficos en las respuestas [...] tuvieron un efecto negativo sobre la humanidad percibida del interlocutor" (sección Resultados, trad. propia), lo que evidencia hasta qué punto el usuario lee el cómo del mensaje, y no solo el qué, como indicio de un interlocutor. El modo en que el *chatbot* construye su discurso, el registro, la coherencia y la reparación del error, es por tanto parte del sistema de signos que produce percepción de consciencia.

Este discurso, además, es hipertextual e hipermedial: el *chatbot* articula texto, enlaces, imágenes, sonidos y tiempos de respuesta en un mismo flujo. Manovich (2021) recuerda que, bajo su diversidad, las interfaces comparten un rasgo: "en ambos casos tratamos con un flujo de datos. Una secuencia de elementos se transmite y se recibe a lo largo del tiempo. [...] Lo que me interesa

es la experiencia del usuario" (sección "Interfaces for a Global Village", trad. propia).

Desde la comunicación, Scolari (2008) sitúa este intercambio en la relación autor, texto, lector, y propone las hipermediaciones como "procesos de intercambio, producción y consumo simbólico que se desarrollan en un entorno caracterizado por una gran cantidad de sujetos, medios y lenguajes interconectados tecnológicamente de manera reticular entre sí" (pp. 113-114).

La convergencia de estos lenguajes en la interfaz es, así, productora de efectos de agencia percibida.

En sexto lugar, se aborda la interfaz como medio y entorno: de McLuhan a las leyes de la interfaz. El autor considera las premisas de McLuhan como un prisma de aplicación de ideas a los nuevos medios, interactuar en estos medios digitales en la sociedad actual conlleva la necesidad de nuevas estrategias ligadas a la interfaz. Todo tiempo pasado sirve como base o experiencia para poder entender potenciales futuros contextos, esto impacta en lo cotidiano en una transformación continua. Los nuevos medios están en constante evolución y el sustento del pasado (medios tradicionales) sirven de plataforma (Logan, 2010).

McLuhan sostiene que el medio es el mensaje, donde, destaca que el papel de las interfaces de los medios no solamente transmiten contenido, sino que además moldean ampliamente la percepción que los seres humanos y las organizaciones sociales tienen. Se denota la importancia de contemplar a las interfaces de los medios y su afectación en el comportamiento en la interacción humana más allá del contenido que éstas posicionan. Los autores destacan que todo nuevo medio se encarga de transformar de manera radical el ritmo y la escala de las interacciones entre humanos, brindando una perspectiva esencialmente útil para el diseño y evaluación de interfaces de usuario (McLuhan, 1964/2003).

Cabe destacar que los medios de comunicación tal como la electricidad y el teléfono proyectan nuestras capacidades sensoriales y cognitivas permitiendo transformar las interfaces con las que interactuamos con el mundo que nos rodea. Ciertas extensiones influencian la manera en la que el mundo articula sus interacciones culturales, sociales, personales y profesionales, con esta idea en

mente resulta fundamental una perspectiva amplia para la mayor comprensión de la interacción entre humanos y sistemas artificiales. (McLuhan, 1964/2003).

Manovich en *la sección* “Interfaces for a Global Village” discute amplias visiones de McLuhan sobre la aldea global se concentran en tiempos de la era digital. Analiza interfaces como blogs y redes sociales, entonces se explora a detalle cómo las diversas plataformas digitales propician una participación más conectada y activa donde buscan la democratización al acceso y divulgación de la información. Sin embargo, destaca limitaciones y desigualdades en el acceso a la información de manera global, sugiriendo que el diseño de las interfaces debe contemplar no tan solo las componentes de funcionalidades sino que también la equidad y accesibilidad a la misma (Manovich, 2021).

En esa línea, Manovich (2021) interpela el diseño mismo de la interfaz: “¿El problema está en nuestros límites cognitivos de procesamiento de información, o en el fracaso de ingenieros y diseñadores para imaginar una interfaz diferente para nuestra aldea global?” (sección “Interfaces for a Global Village”, trad. propia), y recuerda que las interfaces estandarizadas “son solo algunas de las muchas maneras de organizar y presentar texto, transmitir imágenes [...] o datos” (Manovich, 2021, trad. propia).

La interfaz del *chatbot* es, entonces, una elección de diseño entre otras posibles, no una forma natural.

Por otra parte, para comprender el ecosistema y evaluar las interfaces, Scolari (2018) toma en préstamo conceptos y modelos interpretativos de la esfera bionatural y los aplica a la esfera tecnocultural. En *Las leyes de la interfaz*, el autor explora cómo las interfaces no son meras herramientas tecnológicas, sino entidades que existen dentro de un ecosistema cada vez más amplio, en el que los usuarios contribuyen al desarrollo de una tecnología en constante creación, diseño y transformación.

Su enfoque eco-evolutivo, o de transformación basada en el ecosistema, permite apreciar las interfaces no como elementos estáticos orientados a la interacción, sino como participantes activos en un diálogo permanente con los usuarios, su contexto cultural y una realidad tecnológica en expansión. Desde esta perspectiva ecológica, Scolari (2018) amplía la comprensión del modo en

que el escenario tecnológico y los usuarios coevolucionan, y destaca el valor de diseñar interfaces que contemplen las necesidades presentes y futuras de los usuarios, capaces de anticiparse y adaptarse a los cambios.

En séptimo lugar, se aborda la persistencia tecnológica en el diseño de interfaces digitales, apuntando al reconocimiento de la influencia de la cultura y la persistencia tecnológica.

Desde esta mirada se observa si es que subsisten algunos elementos gráficos que están obsoletos, pero que continúan presentes en la interfaz actual del objeto, ya sea porque cumplen la función de recordarle al usuario otra cosa de manera asociativa, porque facilitan el uso de la herramienta, o simplemente porque han quedado predeterminadas. Por ejemplo, asociar el ícono del disquete con la acción de guardar un documento, una caja metálica con dos puntos como la cara de un robot (aunque en el caso de los *chatbots* ni siquiera existen físicamente).

Esta *persistencia* alude al modo en que las herramientas y los nuevos paradigmas tecnológicos, una vez adoptados, se integran de manera duradera en los procesos de diseño y en la interfaz.

El concepto de persistencia no solo remite a la capacidad de la tecnología para evolucionar y adaptarse, sino también al modo en que se heredan y consolidan patrones, funcionalidades y estándares de la experiencia de usuario. Ciertos elementos, principalmente los iconos, se afianzan por su eficacia funcional y por su aceptación, lo que asegura la continuidad de la experiencia y permite que los diseñadores partan de bases ya establecidas sin innovar de manera constante, de modo que se evita la ambigüedad o la confusión y se favorece una experiencia lo más intuitiva posible.

Esta persistencia, según Klinkenberg (2006) puede leerse en clave semiótica como la dimensión diacrónica del código: ciertos signos sobreviven a su función original y significan por su contraste con las formas que los reemplazaron, en los términos de sincronía y diacronía (p. 131).

En octavo lugar, se aborda de manera sucinta la cultura en la interacción digital y en el diseño de interfaces. Se revisa que “La cultura digital se refiere a las prácticas, hábitos y valores que emergen de la interacción humana con las

tecnologías digitales, como internet, redes sociales, aplicaciones móviles, entre otras" (Silva et al., 2023, p. 1).

"En el contexto de la producción y el conocimiento artístico, existe una amplia interacción entre las instancias de cultura material y cultura digital que caracterizan a la cibersociedad" (Costa y de Souza, 2022, p. 1).

Según los autores, la producción y el conocimiento artístico se inscriben en la cibersociedad: surgen de manera continua nuevas necesidades y, con el tiempo, los artistas desarrollan propuestas inéditas para captar la atención de quien contempla sus obras.

"El auge de la globalización gracias al desarrollo de las tecnologías de la información y la comunicación y la expansión del internet trajo consigo el advenimiento de nuevas formas de comunicación e interacciones denominadas la cultura digital" (Prada Nuñez et al., 2022, p. 1).

Como indican los autores, la globalización propicia nuevas formas de interacción en una cibercultura que evoluciona y se transforma de manera cada vez más acelerada.

"Estas formas alternativas de relacionarse y de interactuar entre los grupos sociales representan innovaciones y transformaciones constantes en las distintas esferas de las sociedades, entre ellas la educación" (Prada Nuñez et al., 2022, p. 1).

Según Prada Nuñez et al. (2022), las transformaciones digitales operan como innovaciones en diversas esferas de la sociedad, entre ellas la educación. En ese ámbito, los *chatbots* orientados a lo educativo pueden representar una ventaja significativa para los estudiantes, por ejemplo, al apoyar el estudio de asignaturas específicas o de cursos de idiomas en distintos niveles.

Una característica de la evolución digital es verse afectado por los problemas sociales asociados a las tecnologías de la información y la comunicación, que la ciberdelincuencia puede explotar; los *chatbots* podrían facilitar tareas cada vez más sofisticadas a los ciberdelincuentes. Por ello resulta crucial que el desarrollo de interfaces se asiente sobre una base ética sólida, capaz de prevenir que los futuros agentes operen sin límites en sus respuestas.

Resulta crucial emplear de manera adecuada las redes sociales y las múltiples interfaces disponibles en la vida cotidiana, para que las personas generen contenido de calidad y contribuyan a una cibercultura íntegra, ética y accesible para todas y todos.

En el aula, los estudiantes consultan a los *chatbots* sobre los contenidos que abordan en clase, una práctica que antes realizaban directamente en la web. Como señalan Vivanco y Gorostiaga (2017), los "estudiantes hiperconectados e inmersos en una cultura digital deben dialogar con contenidos culturales que no necesariamente forman parte de su repertorio cultural" (p. 1020).

Al consultar las distintas interfaces disponibles, se busca enriquecer su formación como individuos dispuestos al crecimiento y a la comprensión de la tecnología y la sociedad.

Conviene destacar que las tecnologías de la información y la comunicación constituyen un pilar relevante de la cultura y la interacción digital: "actualmente, las TIC operan como instrumentos culturales que impactan en las formas de percibir, sentir y representar el mundo" (Vivanco y Gorostiaga, 2017, p. 1020).

La influencia de la cultura en el diseño de interfaces de los *chatbots* es un aspecto relevante, reconocido por su eficacia en distintos contextos. La cultura incide en cómo los usuarios perciben la interacción con la tecnología, así como en los estilos de comunicación y los valores que la interfaz transmite.

Esta variabilidad cultural no es externa al análisis semiótico, sino que lo confirma: lo que el usuario reconoce en la interfaz depende de su enciclopedia cultural (Eco, 1976), de modo que un mismo signo, un color, un emoticon, un avatar, activa contenidos distintos según el contexto del usuario.

En noveno lugar, se establece el impacto de la medición de la satisfacción y la relación emocional al usar un *chatbot*. Los usuarios que utilizan *chatbots* potenciados con IA poseen diversas percepciones respecto a la consciencia, llegando incluso hasta en ciertas instancias a pensar que se están comunicando con un ser humano.

Comúnmente en las encuestas es posible levantar datos considerando preguntas acerca de la satisfacción de las respuestas que entrega el *chatbot*, calidad de la interacción y la percepción de la inteligencia del chatbot.

Langarano Guerrero et al. (2022) evaluaron la usabilidad y la satisfacción de un *chatbot* mediante la pregunta "¿Usted sintió que su pedido fue creado de forma satisfactoria a través del chat?" (p. 165), cuyas respuestas posibles eran sí o no.

Los autores reportaron que "un porcentaje mayoritario de usuarios del *chatbot* respondió la encuesta, y de igual forma, la mayoría de los usuarios que la contestaron estuvieron contentos con el desempeño, usabilidad y facilidad de uso del *chatbot*" (Langarano Guerrero et al., 2022, p. 165).

El estudio de Nadarzynski et al. (2019) sobre la aceptabilidad de los servicios de *chatbot* guiados por inteligencia artificial en el ámbito de la salud aborda la cuestión de la empatía y advierte que estos sistemas "no son capaces de manifestar empatía, especialmente para reconocer los estados emocionales de los usuarios" (p. 2, trad. propia).

Pese a ello, algunos sostienen que es posible establecer conexiones emocionales con los *chatbots*: "el uso de señales similares a las de los humanos tuvo una influencia significativa en la conexión emocional" (Araujo, 2018, p. 5, trad. propia). Estos sistemas, además, "tienen un gran potencial en términos de eficiencia del servicio" (Nordheim et al., 2019, p. 317, trad. propia).

> El uso de señales antropomórficas en el diseño y el encuadre de la agencia comunicativa influye en las percepciones de antropomorfismo y de presencia social atribuidas al chatbot; a su vez, el antropomorfismo y la presencia social percibidos inciden en las actitudes hacia la empresa, en la satisfacción del cliente y en el nivel de conexión emocional que las personas sienten con la compañía tras interactuar con el agente (Araujo, 2018, p. 183, trad. propia).

El hallazgo de Araujo (2018) sustenta que el nombre y el estilo lingüístico del *chatbot*, elementos de diseño de bajo costo de implementación, producen efectos perceptuales de alta intensidad. El encuadre de la agencia, es decir, el modo en que se presenta al *chatbot* al inicio de la interacción, modula esos efectos: un *chatbot* presentado como agente inteligente produce mayor presencia social que uno presentado como máquina neutral. Ello sugiere que los *paratextos* de la interfaz, los textos de presentación, los iconos de inicio y los encabezados, son también variables semióticas que la pauta de evaluación debe considerar.

A medida que la tecnología avance, los *chatbots* se integrarán cada vez más en la vida cotidiana, Turkle (2011) reflexiona al respecto:

> Estamos solos pero tememos la intimidad. Las conexiones digitales y el robot sociable pueden ofrecer la ilusión de compañía sin las exigencias de la amistad. Nuestra vida en red nos permite escondernos unos de otros, incluso mientras permanecemos atados unos a otros. Preferimos enviar un mensaje de texto antes que hablar (Turkle, 2011, p. 1, trad. propia).

Estas conexiones con las máquinas impulsarán un diseño que contemple aplicaciones de carácter emocional y que promueva efectos positivos en las personas, en pos de una experiencia más significativa y de mayor aceptación.

2.6.5. Diseño de interacción: affordance, heurística y UX conversacional

El cuarto pilar traduce los conceptos filosóficos y semióticos al lenguaje del diseño, operacionalizándolos en categorías evaluables. Es el pilar más directamente vinculado con la construcción de la pauta.

En primer lugar, dentro del pilar de diseño de interacción, se establece el Affordance, desde Gibson a Norman y continuando de Norman a la interfaz. Gibson (1979) introduce el concepto de *affordance* en la psicología ecológica para designar las posibilidades de acción que un entorno ofrece a un organismo. No son propiedades del objeto ni del sujeto, sino de la relación entre ambos: una silla ofrece la posibilidad de sentarse, pero solo para un organismo con la morfología y el tamaño adecuados.

En palabras de Gibson (1979), “las affordances del entorno son lo que este ofrece al animal, lo que le provee o le suministra, para bien o para mal” (p. 127, trad. propia), y no son “ni una propiedad objetiva ni una propiedad subjetiva; o [son] ambas a la vez” (Gibson, 1979, p. 129, trad. propia).

Norman (1988/2013) reformula el concepto para el diseño de artefactos, distinguiendo entre *affordances reales* y *affordances percibidas*: lo que importa para el diseño no es lo que el objeto puede hacer sino lo que el usuario percibe que puede hacer. Gaver (1991) extiende la tipología: *affordances perceptuales* (visibles al usuario), *affordances ocultas* (existentes pero no visibles) y *affordances falsas* (percibidas por el usuario pero que no corresponden a funciones reales).

Norman (2013) precisa el carácter relacional del concepto: “una affordance es una relación entre las propiedades de un objeto y las capacidades del agente que determinan cómo podría usarse el objeto” (p. 11, trad. propia), y advierte que “en el diseño, los significantes son más importantes que las affordances, pues comunican cómo usar el diseño” (Norman, 2013, p. 19, trad. propia).

Gaver (1991) complementa lo anterior al recordar que “cuando las affordances son perceptibles, ofrecen un vínculo directo entre percepción y acción; las affordances ocultas y falsas conducen a errores” (p. 79, trad. propia).

En la interfaz del *chatbot*, los significantes, el indicador de escritura, el avatar, el nombre, son los que comunican al usuario que hay un agente con el cual interactuar.

Aplicado a la interfaz de *chatbot*, el *affordance* permite analizar qué elementos de diseño invitan al usuario a percibir agencia, intencionalidad e identidad en el sistema. Un avatar animado que parpadea ofrece un *affordance* de presencia viva; el punto de escritura (typing indicator) ofrece un *affordance* de pensamiento en curso; un nombre propio para el *chatbot* ofrece un affordance de identidad individual. Estos *affordances* no producen automáticamente la percepción de consciencia, pero son sus condiciones de posibilidad desde el diseño.

A la lectura del affordance conviene sumar un criterio de simplicidad. Maeda (2006) sostiene que “la forma más simple de lograr la simplicidad es a través de una reducción reflexiva. En caso de duda, elimina. Pero ten cuidado con lo que eliminas” (p. 1, trad. propia) y que “la simplicidad consiste en sustraer lo obvio y agregar lo significativo” (Maeda, 2006, p. ix, trad. propia).

Para esta tesis, esto implica que la percepción de consciencia no requiere acumular señales, sino seleccionar las significativas. Maeda (2019) añade una clave epistemológica al describir la computación como “un universo invisible y ajeno [...], una especie de materia prima que no obedece las leyes de la física” (p. xi, trad. propia), lo que justifica tratar al *chatbot* no como un objeto fijo, sino como un sistema en transformación cuya lógica el diseño debe hacer perceptible al usuario.

Gnewuch et al. (2017) proponen principios de diseño orientados a producir cooperación social, entre ellos la transparencia sobre las capacidades del sistema, la consistencia del tono y la personalidad, y la capacidad de reparar la conversación cuando ocurren malentendidos.

Para esta investigación, los principios abordados por Gnewuch et al. (2017) son antecedentes operacionales de los criterios de la pauta: cada uno de ellos puede reinterpretarse como una condición de posibilidad para la percepción de consciencia. Un sistema que es transparente sobre sus limitaciones pero las comunica con expresiones que simulan consciencia de sí mismo produce más atribución de agencia que uno que emite mensajes de error genéricos.

De acuerdo con Gnewuch et al. (2018) los estados del sistema y su efecto en la percepción de deliberación y presencia demuestran que el efecto del indicador de escritura con una demora dinámica en la respuesta del chatbot, simulada por el indicador de escritura en curso, causa mayor percepción de presencia social que una respuesta inmediata, siempre que esa demora sea percibida como natural y proporcional al contenido de la respuesta. Una respuesta instantánea puede parecer mecánica; una demora excesiva produce frustración. La zona de percepción de naturalidad es estrecha y culturalmente variable.

Lo anterior constituye el sustento empírico directo a la subdimensión de fluidez temporal que esta tesis propone como parte del Pilar 4 (Dis. 4b, affordance temporal de Gaver): el ritmo y la latencia percibidos como naturales son indicadores evaluables de la percepción de consciencia.

Gnewuch et al. (2022) sostienen que “las personas tienen expectativas sobre el comportamiento no verbal de humanos y no humanos” (p. 776, trad. propia). Por ello, el diseño debe contemplar señales semióticas que propicien una recepción positiva por parte del usuario: la teoría de la violación de expectativas resulta también aplicable a los *chatbots*.

Følstad y Brandtzaeg (2017) sitúan a los *chatbots* ante un cambio de paradigma: “El diseño para *chatbots* representa una transición desde el diseño de la disposición visual y los mecanismos de interacción hacia el diseño de la conversación” (p. 38, trad. propia).

Las conclusiones de Følstad y Brandtzaeg (2017) son especialmente pertinente para esta tesis porque establece la tensión que el diseño de interacción debe resolver: si la conversación desplaza a los elementos visuales como canal principal, ¿qué rol cumplen el avatar, la tipografía, el color y los indicadores de escritura? Esta investigación argumenta que esos elementos no son residuales sino constitutivos de la experiencia, y que su análisis semiótico es indispensable para comprender la percepción de consciencia.

Piccolo et al. (2019) advierten que “un entusiasmo potencial de los usuarios puede derivar en frustración o decepción, y en rechazo, si la tecnología no satisface sus expectativas” (sección Introducción, trad. propia), y que predomina aún “los desafíos técnicos por sobre la comprensión de las percepciones, expectativas y contextos de uso de los usuarios” (Piccolo et al., 2019, Resumen, trad. propia).

El hallazgo central de Piccolo et al. (2019) es directamente pertinente para la vacancia que esta tesis doctoral busca cubrir: los atributos humanizados de los *chatbots* requieren un enfoque multidisciplinario para su detección y evaluación, y ese enfoque aún no está abordado completamente. Piccolo et al. (2019) señalaron también que la brecha entre el diseño del sistema y las expectativas del usuario era el principal obstáculo de la interacción, lo que anticipa la hipótesis de esta investigación: los signos de la interfaz producen expectativas de presencia y agencia que el sistema puede o no sostener.

Para esta investigación, Gambino et al. (2020) realizan una extensión del paradigma CASA (*Computers Are Social Actors*) tiene consecuencias directas: significa que la percepción de consciencia en el *chatbot* no solo depende de señales de semejanza humana, sino también de la coherencia del sistema con las expectativas que el usuario tiene de los agentes conversacionales como clase de objeto cultural. El diseño de la interfaz debe, por tanto, operar en dos registros simultáneos: el de la humanización (que activa los guiones humano-humano) y el de la legibilidad como *chatbot* (que activa los guiones humano-medio), y la pauta de evaluación de esta tesis debe poder capturar ambas dimensiones.

Lo expresan así: “los humanos pueden desarrollar y aplicar guiones sociales humano-medio a esas interacciones” (Gambino et al., 2020, p. 71, trad. propia).

En relación con el diseño de interfaz, autores como Martín-Barbero, McLuhan, Manovich y Scolari abordan la interfaz, las interacciones digitales y la experiencia de usuario en un sentido amplio que conviene trasladar al estudio de los *chatbots*. Scolari (2018) entiende las interfaces como entornos de interacción en los que distintos actores, humanos y tecnológicos, intercambian información y ejecutan acciones, negocian, discuten, acuerdan o rompen relaciones; mientras una herramienta se utiliza, un entorno se habita (p. 30).

Scolari (2018) señala que una interfaz bien diseñada se vuelve transparente y permite al usuario concentrarse en la tarea que realiza. Asimismo, considera el diseño de información y la usabilidad como aspectos fundamentales para lograr una interfaz bien resuelta. En esa línea, el *chatbot* constituye un servicio virtual que requiere precisamente una interfaz con tales características.

Con respecto al diseño de la interfaz, uno de los desafíos consiste en mejorar las experiencias en el ámbito digital aprovechando los datos del comportamiento del usuario. El *big data* designa un gran volumen de información digital que, no obstante, tiene límites: las experiencias diseñadas determinan el tipo de datos que pueden extraerse de ellas. Lo que estos datos permiten es informar sobre el modo en que los usuarios disfrutan de un producto o servicio, y de allí pueden derivarse enseñanzas y principios de diseño aplicables con posterioridad.

Una vez que se comprende de manera anticipada qué buscarán los usuarios, bajo el paradigma del procesamiento inductivo del *big data*, la inteligencia artificial puede prever sus necesidades y favorecer un diseño estratégico capaz de ofrecer el contenido o la interacción adecuados en el momento oportuno. Bien analizado e implementado, esto mejora la efectividad y permite satisfacer las expectativas del usuario al utilizar una aplicación.

De Souza et al. (2021) remiten el término avatar a su origen religioso y sánscrito *avatara*, que designaba el descenso de una deidad a la tierra y las formas humanas o animales que asumía el dios hindú Vishnú al visitar a la

humanidad. El concepto se popularizó en el mundo digital a partir de la novela *Snow Crash*, de Stephenson (1992), que describe un entorno virtual tridimensional en el que los usuarios pueden adoptar personalidades o habitar otras formas de existencia. Este imaginario, heredado de la literatura y el cine, contribuye a que algunos usuarios proyecten en los *chatbots* atributos de identidad e incluso de consciencia.

En el escenario habitual de interacción humano-máquina mediante un *chatbot* ocurren situaciones como la solicitud de asistencia de un usuario a un asistente virtual, sea porque es el canal más inmediato que la empresa o la institución pone a disposición, sea porque el usuario prefiere no comunicarse telefónicamente con un operador humano que, tras un protocolo corporativo y la selección de múltiples opciones, suele entregar una respuesta estandarizada, cargada de información técnica y sin registro escrito para una consulta posterior.

El fin último de estas gestiones es obtener información, resolver un problema o contratar un producto o servicio de manera remota. A ello también puede responder un bot que proyecta, en un acto de comunicación sostenido por la interfaz, una imagen digital cuyo signo de identidad más representativo es el avatar, personalizado por lo común en coherencia con los valores corporativos. El bot emplea como canales de transmisión los mensajes de texto, los sonidos, los audios, los hiperenlaces y las imágenes desplegadas en el chat en milésimas de segundo.

El discurso textual lo configura un sistema informático automatizado que recurre al *big data* para acceder a la información alojada en la red, consulta el historial para detectar si la dirección IP del dispositivo se conectó con anterioridad y anticipar la intención del usuario, y procesa los datos del dispositivo en el momento del contacto para entregar una respuesta personalizada.

De este modo, el bot emplea códigos del lenguaje comprensibles para el usuario: utiliza la geolocalización para ajustar el idioma local y revisa la configuración del dispositivo para identificar opciones de accesibilidad, de manera que, si el usuario tiene baja visión o es ciego, la respuesta prevista mediante textos e imágenes muta a una comunicación basada en audios y en la recepción del *feedback* por voz y desplazamientos táctiles sobre la pantalla. Esto se

refuerza con la mirada de Revilla y Carreras (2025) "La accesibilidad web se define como el conjunto de características que un sitio web, aplicación móvil o documento digital debe incorporar para que el mayor número posible de personas, en el mayor número posible de circunstancias, pueda percibirlo, usarlo y comprenderlo." (p. 5, traducción propia).

El diseño de interfaces de usuario constituye un pilar fundamental de la interacción humano-computador en contextos que vinculan a los *chatbots* con aplicaciones tecnológicas diversas.

El Diseño Centrado en el Usuario es un enfoque que orienta la creación de interfaces efectivas y accesibles. Resulta crucial que la presentación de la información facilite la interacción y la experiencia del usuario, que puede variar según la usabilidad y según la percepción de consciencia atribuida a los *chatbots*.

Conceptualmente, la interfaz visible tiende a ser percibida por el usuario como la aplicación misma, pues es la parte con la que interactúa directamente; sin embargo, su usabilidad depende tanto del diseño visual como de la arquitectura, estructura y organización del sistema (Cobarrubias Soto et al., 2022, p. 79).

Su énfasis recae en la comprensión de las necesidades de los usuarios, pues exige un conocimiento profundo de su comportamiento para crear interfaces que faciliten un aprendizaje sencillo e intuitivo orientado a su satisfacción.

El proceso de diseño de la interfaz se organiza en etapas, desde la exploración o investigación inicial hasta la evaluación y la participación del usuario. Díaz et al. (2013) presentan un enfoque sistemático que considera la evolución de la interacción entre computadores y seres humanos, junto con sus implicaciones de evaluación, como las métricas y el ciclo de vida de la interfaz.

Diseñar una interfaz de usuario es, así, un proceso de etapas definidas en el que diseñadores y desarrolladores iteran de manera continua a partir de la retroalimentación de los usuarios, para asegurar que la interfaz materialice sus expectativas y necesidades.

Las herramientas de diseño y prototipado resultan relevantes para desarrollar interfaces de usuario efectivas. Vargas Márquez et al. (2021) señalan que el *Design Thinking*, aplicado a la experiencia de usuario, subraya la

importancia de desarrollar empatía con las personas para comprender sus necesidades reales.

El trabajo con prototipos permite evaluarlos y mejorarlos a partir de la retroalimentación de los usuarios, cuyo aporte resulta esencial para un desarrollo más efectivo y centrado en las personas.

En el contexto de los *chatbots*, este enfoque puede orientar la implementación de interfaces conversacionales que buscan emular la interacción humana y que, al hacerlo, pueden incrementar la percepción de empatía y de consciencia en el intercambio.

Un componente crítico del diseño de interfaces es la evaluación de la usabilidad. González Hernández y Victoria-Uribe (2023) ofrecen una perspectiva que recurre a métodos cualitativos, como los perfiles de usuario y las pruebas de usabilidad, para mejorar la experiencia en entornos de realidad virtual.

Esta evaluación permite que los diseñadores identifiquen tempranamente problemas de usabilidad y realicen los ajustes necesarios para mejorar tanto la usabilidad como la satisfacción del usuario.

Las implicaciones éticas y tecnológicas del diseño de interfaces en sistemas artificiales como los *chatbots*, cada vez más sofisticados, pueden incidir en la percepción de consciencia e inteligencia que se atribuye a estos sistemas.

La intencionalidad con que se presenta la información facilita la interacción del usuario con el sistema, por lo que resulta crucial adoptar un enfoque ético en el desarrollo de interfaces que respete la privacidad y la autonomía de las personas.

En síntesis, el diseño de interfaces de usuario resulta crítico para la interacción, en especial en contextos donde los *chatbots* se potencian con inteligencia artificial, pues la aplicación del diseño centrado en el usuario, el prototipado y la evaluación de la usabilidad es fundamental para crear interfaces efectivas y accesibles.

Por otra parte, la tecnología se transforma de manera continua, de modo que el diseñador debe contemplar la dimensión ética del desarrollo tecnológico, ya que sus decisiones impactan de forma masiva en los usuarios finales.

Garantizar interfaces funcionales que fomenten una interacción positiva y respetuosa con las personas es una tarea central.

En cuanto a la experiencia de usuario ante agentes percibidos como conscientes, algunos usuarios reportan confusión respecto de si interactúan con un humano o con una inteligencia artificial, e incluso se preguntan por la veracidad de sus respuestas. De cara a un eventual estadio de inteligencia artificial general (AGI), cabe preguntarse si el comportamiento de estos sistemas se modificará a medida que se vuelvan más sofisticados y si conservarán la disposición servicial y amable de los primeros agentes.

Dado que la literatura no detalla cómo se integraría la consciencia en la inteligencia artificial, conviene considerar que muchas interacciones cotidianas, como una llamada telefónica, pueden ser realizadas por un sistema basado en IA. En tales casos, el *chatbot* está preparado para interactuar en tiempo real con un ser humano que, con frecuencia, no logra discernir si conversa con una persona o con un sistema. Esa delgada línea interpela de manera directa a la ética de la interacción.

Aunque la literatura no profundiza ni alcanza consenso sobre la posibilidad de que algunos *chatbots* lleguen a tener una consciencia limitada para aspectos específicos de su diseño, el actual auge de la inteligencia artificial ha dado lugar tanto a agentes que ejecutan tareas de manera automatizada como a agentes que se autoensamblan según las necesidades de sus diseñadores.

En el último tiempo, algunas empresas han habilitado funciones en fase *beta* que integran sus agentes con el escritorio del computador para ejecutar tareas avanzadas según los requerimientos del usuario. Resulta notable que los avances en las interfaces y la creciente sofisticación de los *chatbots* alcancen, en el marco de la Agenda 2030, límites difíciles de imaginar pocos años atrás.

El concepto de consciencia aplicado a la inteligencia artificial aún no se aborda desde la perspectiva de la ejecución de tareas específicas. Existe, por tanto, la posibilidad de que la consciencia atribuida a un *chatbot* se constituya en objeto de estudio para los teóricos de la ética, la filosofía y la tecnología.

Para esta investigación, la contribución de Chen et al. (2024) es metodológica y conceptual: muestra que las señales antropomórficas de la

interfaz influyen en la atención visual y, junto con el estilo conversacional, en la presencia social percibida, la confianza y la satisfacción. Estos resultados aportan un antecedente empírico para examinar los elementos gráficos como componentes relevantes de la experiencia del usuario y fundamentan la pertinencia de su análisis semiótico sistemático, sin establecer una relación directa con la percepción de consciencia o la atribución de intencionalidad.

En segundo lugar, se abordan las heurísticas de Nielsen y su reinterpretación para el análisis de la percepción de consciencia. El fundamento se recupera de *Usability engineering* (Nielsen, 1994), mientras que la presentación de los diez principios utiliza las denominaciones y descripciones de la versión web revisada en 2024 (Nielsen, 2024). Estos principios, orientados originalmente a la usabilidad de interfaces, se reinterpretan en esta investigación como criterios para examinar condiciones de diseño que pueden favorecer o limitar la atribución de consciencia al *chatbot*. Según Nielsen (2024), los diez principios heurísticos de usabilidad son:

- Principio Nº 1: Visibilidad del estado del sistema: la interfaz debe informar al usuario qué está ocurriendo, mediante retroalimentación clara y oportuna.
- Principio Nº 2: Correspondencia entre el sistema y el mundo real: el sistema debe usar lenguaje, conceptos, metáforas e información comprensibles para el usuario, no solo términos técnicos.
- Principio Nº 3: Control y libertad del usuario: el usuario debe poder deshacer, rehacer, cancelar acciones o salir fácilmente de estados no deseados.
- Principio Nº 4: Consistencia y estándares: la interfaz debe mantener coherencia visual, funcional y terminológica, respetando convenciones conocidas por los usuarios.
- Principio Nº 5: Prevención de errores: es mejor evitar que el error ocurra antes que limitarse a mostrar mensajes después del problema.
- Principio Nº 6: Reconocimiento antes que recuerdo: la interfaz debe reducir la carga de memoria del usuario, mostrando opciones, instrucciones o información contextual cuando sea necesario.

- Principio Nº 7: Flexibilidad y eficiencia de uso: el sistema debe servir tanto a usuarios principiantes como avanzados, permitiendo atajos, personalización o formas más rápidas de interacción.
- Principio Nº 8: Diseño estético y minimalista: la interfaz no debe sobrecargar al usuario con información irrelevante o innecesaria.
- Principio Nº 9: Ayudar a reconocer, diagnosticar y recuperarse de errores: los mensajes de error deben ser claros, explicar el problema y sugerir una solución.
- Principio Nº 10: Ayuda y documentación: aunque lo ideal es que el sistema sea autoexplicativo, debe existir ayuda accesible, clara y orientada a tareas concretas.

Por ejemplo, la heurística de visibilidad del estado del sistema, “el diseño debería mantener siempre informados a los usuarios sobre lo que ocurre, mediante retroalimentación apropiada en un tiempo razonable” (Nielsen, 2024, sección “1: Visibility of System Status”, párr. 1, traducción propia), se reinterpreta aquí como la base del indicador de escritura que produce la percepción de un agente que “piensa”. La correspondencia con el mundo real, “el diseño debería hablar el lenguaje de los usuarios” (Nielsen, 2024, sección “2: Match Between the System and the Real World”, párr. 1, traducción propia), sostiene la atribución de un interlocutor que comprende. Y la consistencia y los estándares fundan la coherencia identitaria del agente a lo largo de la conversación.

A las heurísticas de Nielsen se suman los principios o reglas de Shneiderman y Plaisant (2005), presentados de manera explícita como saber acumulado y adaptable: “Estos principios, derivados de la experiencia y refinados durante dos décadas, requieren validación y ajuste para dominios de diseño específicos” (p. 75, trad. propia). Constituyen un conjunto de principios orientadores para evaluar la usabilidad de sistemas interactivos. Entre ellos, “buscar la consistencia” y “permitir la reversión fácil de las acciones” (Shneiderman y Plaisant, 2005, p. 75, trad. propia), pueden ser directamente reinterpretables como condiciones que sostienen o erosionan la identidad percibida del agente conversacional.

Las ocho reglas de oro del diseño de interfaces de Shneiderman y Plaisant (2005), presentadas en su libro *Designing the User Interface*, son:

- Regla Nº 1: Buscar la consistencia: usar secuencias de acciones, terminología, menús, colores y tipografía coherentes en toda la interfaz. La consistencia reduce la carga cognitiva y facilita el aprendizaje.
- Regla Nº 2: Permitir el uso de atajos (*shortcuts*): los usuarios frecuentes necesitan realizar tareas más rápidamente. Abreviaciones, teclas de función, comandos ocultos y macros son mecanismos que responden a esta necesidad.
- Regla Nº 3: Ofrecer retroalimentación informativa: cada acción del usuario debe tener una respuesta visible del sistema. Las acciones frecuentes o menores requieren respuestas modestas; las poco frecuentes o mayores requieren respuestas más sustantivas.
- Regla Nº 4: Diseñar diálogos que indiquen cierre: las secuencias de acciones deben organizarse en grupos con un inicio, una parte intermedia y un final. La retroalimentación al completar un grupo de acciones da al usuario la satisfacción de haber logrado algo.
- Regla Nº 5: Prevenir errores y facilitar su manejo: el sistema debe diseñarse para que los usuarios no puedan cometer errores graves. Si ocurre un error, el sistema debe detectarlo y ofrecer instrucciones simples para recuperarse.
- Regla Nº 6: Permitir la reversión fácil de acciones (*undo*): esta característica alivia la ansiedad, ya que los usuarios saben que los errores pueden deshacerse. Fomenta la exploración de funciones desconocidas.
- Regla Nº 7: Apoyar el sentido de control interno (*locus of control*): los usuarios experimentados quieren sentir que controlan la interfaz y que esta responde a sus acciones. Deben ser ellos quienes inicien las acciones, no reaccionar ante ellas.
- Regla Nº 8: Reducir la carga de memoria a corto plazo: las limitaciones de procesamiento de información humana requieren que las pantallas sean simples, que las páginas múltiples se consoliden, y que los movimientos de ventanas sean reducidos. Para ello se recomienda disponer de suficiente tiempo de entrenamiento, ayudas en pantalla y resúmenes accesibles.

Los llamados principios de oro son precisamente las ocho reglas de oro del diseño de interfaces o *Eight Golden Rules of Interface Design*. Históricamente se atribuyen a Ben Shneiderman, pero en la edición de 2005 del libro *Designing the User Interface: Strategies for Effective Human-Computer Interaction* aparece en una obra coautorada por Shneiderman y Catherine Plaisant.

Más adelante, los aportes de Nielsen (1994), presentados con las denominaciones de las heurísticas recogidas en Nielsen (2024), serán comparados con las ocho reglas de oro de Shneiderman y reinterpretados para el análisis de la percepción de consciencia en *chatbots*, con el propósito de complementar la dimensión evaluativa de la pauta.

En tercer lugar, se aborda la UX conversacional y diseño de la interacción con *chatbots*. Las guías de diseño conversacional de Google (2024) y Microsoft (2023) establecen principios específicos para el diseño de *chatbots* que son relevantes para la construcción de la pauta. Conceptos como *turn-taking* (gestión de los turnos conversacionales), personalidad del asistente, naturalidad y expectativas del usuario, retroalimentación conversacional (Google), o claridad de intención, transparencia del sistema, manejo del error y coherencia de identidad (Microsoft) proveen un vocabulario técnico que complementa el marco filosófico y semiótico.

Microsoft (2023) resume el desplazamiento que define la experiencia conversacional: “Otras modalidades cargan al usuario con la tarea de aprender comportamientos de interacción significativos para el sistema [...]. La CUX da vuelta la situación: en lugar de que los usuarios tengan que aprender el sistema, es el sistema el que aprende” (sección “Conversational user experience”, trad. propia). Google (2024), por su parte, recuerda que “al tomar turnos evitamos interrumpirnos y mantenemos la conversación sincronizada” (sección “Turn-taking”, trad. propia), de modo que la gestión del turno es un signo de competencia conversacional que el usuario atribuye al agente.

La UX conversacional no es solo un conjunto de buenas prácticas: es un campo que ha construido un conocimiento empírico sobre qué condiciones de diseño producen experiencias de interacción satisfactorias. Al reinterpretarlo desde el marco de esta investigación, este conocimiento permite identificar qué

condiciones de satisfacción se superponen con, o son condición de, la percepción de consciencia.

La UX conversacional se centra en el diseño de la experiencia de usuario para interfaces conversacionales, con el propósito de anticipar las necesidades del usuario y crear interacciones fluidas y naturales. Comprende, además, el estudio de los modelos mentales del usuario y de su contexto de uso.

Finalmente, las WCAG 2.2 (W3C, 2023) aportan cuatro principios, perceptibilidad, operabilidad, comprensibilidad y robustez, que esta investigación reinterpreta como criterios e indicadores de la pauta: la interfaz debe presentarse "de modo que [los usuarios] puedan percibirlos", ser "operable", "comprensible" y "suficientemente robusto" (W3C, 2023, trad. propia). La accesibilidad deja de ser un requisito externo y se integra como dimensión del análisis semiótico e interactivo.

El diseño de la conversación se centra en la creación de diálogos efectivos entre el chatbot y el usuario, de modo que las interacciones resulten satisfactorias y coherentes. Implica definir los flujos conversacionales y adaptar el lenguaje al contexto del usuario.

El método de evaluación de comunicabilidad (MEC) propuesto por de Souza (2005) evalúa si el usuario comprende el mensaje de diseño codificado en la interfaz, identificando rupturas comunicacionales mediante el análisis de transcripciones de interacción. Para esta investigación, el MEC es el antecedente metodológico más cercano a la pauta que esta tesis propone: ambos son instrumentos cualitativos de evaluación de interfaces con base semiótica.

De Souza (2005) define la unidad de análisis de la ingeniería semiótica en términos directamente articulables con este marco:

> un sistema interactivo basado en computadora, diseñado para comunicar a sus usuarios la lógica que lo sustenta y los principios interactivos mediante los cuales el propio artefacto puede ser utilizado para ayudar a los usuarios a alcanzar una determinada gama de objetivos (p. 29, traducción propia).

La diferencia con esta tesis es de objetivo: el MEC evalúa la comunicabilidad funcional, mientras que la pauta propuesta evalúa la producción de percepción de consciencia.

Para de Souza (2005) la diferencia es que el MEC evalúa comunicabilidad funcional, mientras que la pauta de esta tesis evalúa producción de percepción de consciencia. Esa diferencia define el aporte de esta investigación frente al trabajo fundacional desarrollado por de Souza y sus colaboradores.

En la investigación de Westerman et al. (2019), el efecto de los errores tipográficos sobre la humanidad percibida tiene implicaciones para el análisis del discurso conversacional como dimensión semiótica: el modo en que el *chatbot* construye sus respuestas, el registro, la coherencia y la gestión del error son aspectos que esta tesis propone examinar en relación con la percepción de consciencia, más allá de los elementos puramente visuales.

Borsci et al. (2022) proponen la *Bot Usability Scale* (BUS-15), un cuestionario de 15 ítems que evalúa la satisfacción con la experiencia de interacción con *chatbots*. Comprende cinco factores: accesibilidad percibida, calidad percibida de las funciones del agente, calidad de la conversación y de la información proporcionada, privacidad y seguridad percibidas, y tiempo de respuesta. Sus ítems incluyen aspectos como la continuidad conversacional, el seguimiento del contexto y el manejo de situaciones en las que la conversación resulta poco clara, ampliando la evaluación hacia particularidades de los agentes conversacionales.

Para esta investigación, BUS-15 constituye un antecedente evaluativo cercano a la pauta propuesta y evidencia la necesidad de herramientas específicas para los *chatbots*. Sin embargo, su propósito es evaluar la satisfacción del usuario con la interacción y la calidad percibida del servicio, no examinar específicamente las condiciones de diseño vinculadas con la atribución de consciencia o agencia. Aunque incluye ítems sobre comprensión y seguimiento del contexto, estos se interpretan en relación con la experiencia de uso. La presente tesis retoma ese antecedente y desplaza el foco hacia la articulación semiótica, fenomenológica e interactiva de dicha atribución.

Cheng et al. (2022) distinguen dos regímenes de relación que el usuario puede establecer con un chatbot: las normas de intercambio (transaccionales, centradas en la eficiencia) y las normas comunales (relacionales, centradas en el cuidado mutuo). Los autores las definen así: "una relación de intercambio se rige

por el intercambio mutuo de beneficios comparables", mientras que "una relación comunal se basa en aspectos sociales, como atender las necesidades del otro" (Cheng et al., 2022, p. 2, trad. propia); de ahí que "las relaciones de intercambio suelan considerarse más impersonales que las comunales, que ponen mayor énfasis en la amistad" (Cheng et al., 2022, p. 2, trad. propia). Respecto del efecto sobre la confianza, concluyen que "las relaciones de intercambio refuerzan la importancia de la competencia percibida sobre la confianza, mientras que las relaciones comunales no moderan el efecto de la calidez percibida sobre la confianza" (Cheng et al., 2022, p. 1, trad. propia).

Para esta investigación, la distinción entre normas relacionales y transaccionales que realizan Cheng et al. (2022) es un criterio evaluativo que la pauta debe incluir: el diseño visual e identitario del chatbot activa uno u otro régimen de normas, y esa activación tiene consecuencias medibles en la satisfacción y la confianza del usuario.

2.6.6. Síntesis del marco teórico para la construcción de la pauta metodológica

El marco teórico desarrollado articula cuatro tradiciones en una sola respuesta disciplinar a la vacancia identificada. El Pilar 1 establece la condición epistemológica que hace posible el objeto: la percepción de consciencia puede estudiarse sin afirmar que el sistema sea consciente, porque el problema se desplaza de la ontología de la máquina a la experiencia del usuario que la percibe. El Pilar 2 describe esa experiencia desde la fenomenología y el enactivismo: la consciencia percibida no es un error del usuario, sino una forma de acoplamiento situado que se produce en el acto mismo de la interacción. El Pilar 3 aporta el instrumental para analizar las señales que sostienen esa atribución, al leer los elementos visuales, verbales e hipermediales de la interfaz como un código que produce significación. El Pilar 4 traduce todo lo anterior al lenguaje del diseño, operacionalizando *affordances*, heurísticas y criterios de experiencia conversacional en condiciones evaluables.

Leídos en conjunto, los cuatro pilares describen un mismo fenómeno en cuatro niveles complementarios: por qué es legítimo estudiarlo, cómo se vive, con qué signos se produce y mediante qué decisiones de diseño puede orientarse. La

consciencia percibida queda definida, en síntesis, como un fenómeno semiótico emergente y relacional: emergente porque no preexiste a la interacción, y relacional porque resulta del encuentro entre los signos de la interfaz y la enciclopedia cultural y la disposición fenomenológica del usuario. Esta definición es la que permite a la investigación escapar de la obsolescencia tecnológica, ya que lo que cambia con cada generación de sistemas es la tecnología, no la lógica de la semiosis ni las condiciones interactivas de la experiencia.

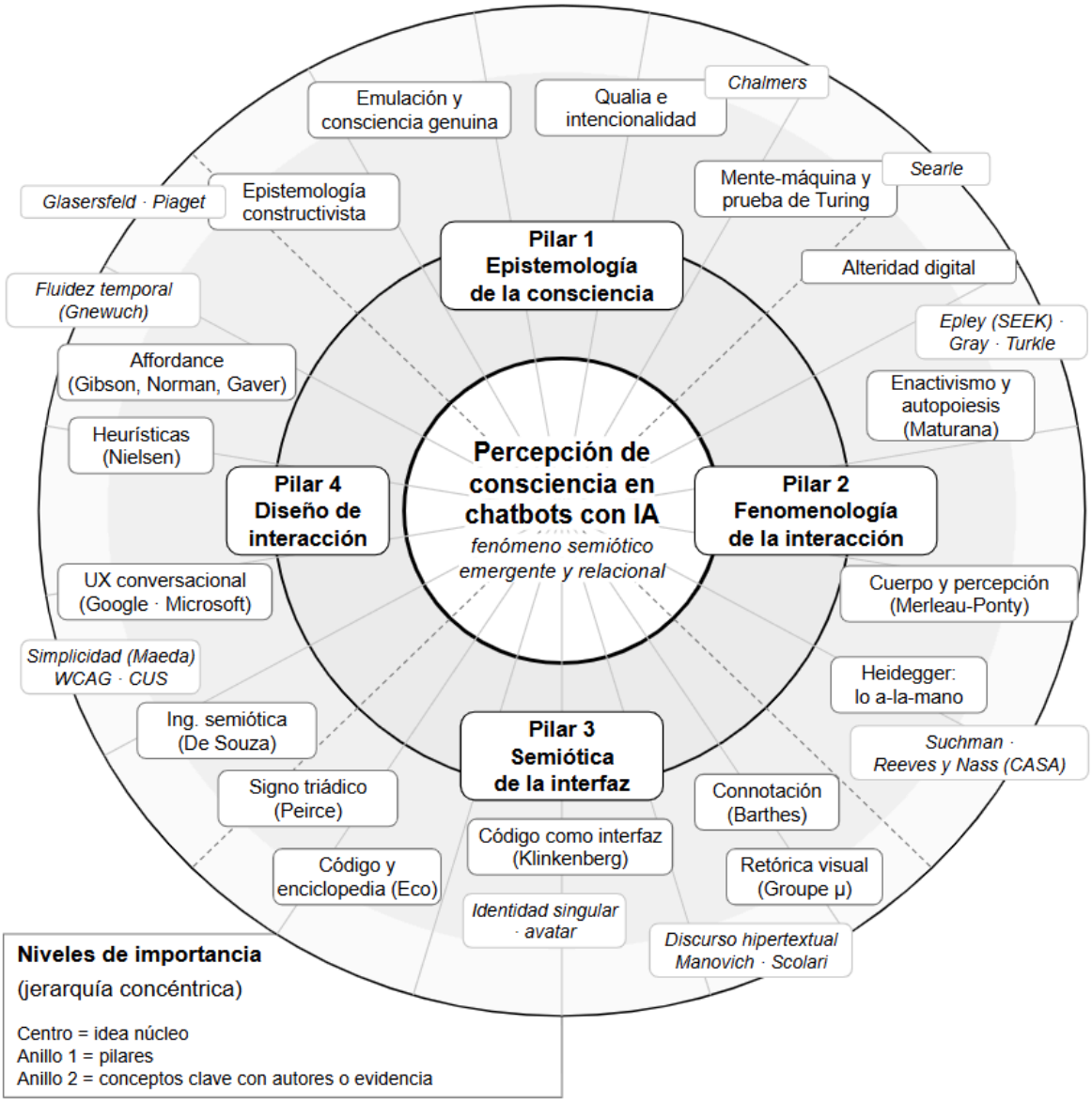
De esta articulación se desprenden los conceptos operacionales que la pauta de evaluación traducirá en criterios e indicadores: la identidad singular del agente, la alteridad digital, el *affordance* de presencia, agencia e intencionalidad, la fluidez temporal de la respuesta, la connotación visual del avatar y la coherencia entre las señales de diseño y la naturaleza del sistema. En la figura 1 se puede observar la articulación conceptual.

En la siguiente página, se observa la figura 1 que presenta la articulación conceptual, los conceptos operacionales remite a uno o varios pilares, manteniendo la consistencia con las preguntas, hipótesis y objetivos, de modo que el marco teórico no es un prefacio erudito, sino la base argumentativa directa del instrumento que esta tesis se propone construir.

Con estos fundamentos definidos se da término al marco teórico y en el apartado siguiente, que es el marco metodológico, los conceptos se operacionalizan en categorías de acuerdo al análisis documental de las fuentes y se ordenan en la matriz que después de completar las fases del proceso de investigación se dará origen a la pauta de evaluación metodológica de la percepción de consciencia en *chatbots* inteligentes.

Figura 1

Mapa conceptual del marco teórico



Nota. Elaboración propia a partir del marco teórico de esta investigación.

2.7. Marco Metodológico

2.7.1. Enfoque y tipo de investigación

La investigación adopta un enfoque cualitativo e interpretativo, con orientación aplicada al diseño por medio de un resultado propositivo: su producto principal no es sólo la descripción e interpretación de un fenómeno empírico particular, sino la construcción de un instrumento conceptual transferible, la pauta metodológica de evaluación, que otros investigadores y profesionales del diseño podrán aplicar y adaptar. En este sentido, la investigación se inscribe en la tradición de la investigación en diseño orientada a la producción de conocimiento metodológico, comparable con trabajos como los de Frascara (2004) y Cross (2007).

Frascara (2004) propone una visión del diseño de comunicación centrada en la resolución de problemas a través de métodos sistemáticos y replicables, entendiendo al diseñador como un profesional que produce conocimiento aplicado mediante principios y procedimientos que pueden transferirse a distintos contextos comunicacionales, más allá del caso particular que los origina.

Por su parte, Cross (2007) caracteriza al diseño como una forma de conocimiento propia, distinta de la científica y la humanística, cuyo modo de indagación produce un tipo de saber que se materializa en procesos, métodos y artefactos conceptuales transferibles, validando la investigación en diseño como una vía legítima para generar conocimiento metodológico y no solo soluciones puntuales.

En este sentido, ambos autores respaldan la idea de que la investigación en diseño puede tener como resultado legítimo un instrumento o método transferible, y no únicamente la interpretación de un caso particular. Esta perspectiva resulta coherente con el carácter propositivo de la presente tesis, cuyo aporte principal consiste en la construcción de una pauta de evaluación metodológica aplicable más allá del corpus específico analizado. Dicha pauta opera como una matriz maestra transferible, orientada a fundamentar futuras evaluaciones contextualizadas, y no como una rúbrica cerrada, una lista de cotejo ni una escala psicométrica.

Sobre las técnicas necesarias para llevar a cabo esta investigación, se encuentran el análisis de información documental, análisis de contenido y discurso, observación participante, ejecutando diversos procedimientos y tareas, y generando un registro de respuestas para obtener los datos que serán analizados e interpretados. Las entrevistas a expertos permiten reafirmar orientaciones o abordar los tópicos que no fueron aclarados en las etapas anteriores de la investigación, facilitando la comprobación o refutación de hipótesis y las conclusiones.

Los entrevistados, seleccionados por pertinencia disciplinar, corresponden principalmente a profesionales y doctores asociados a una misma universidad pública, la Universidad de Santiago de Chile. En un principio la formulación inicial fue la de selección por conveniencia, sin embargo se reemplazó por pertinencia disciplinar ya que los participantes cuentan con formación en distintas universidades y experticia en distintos campos requeridos, pero al mismo tiempo mantienen la cercanía sociocultural necesaria para lograr comprender e interpretar el fenómeno en un contexto más cercano al que se inscribe esta investigación para facilitar el análisis de los resultados.

En cuanto a tipo y grado de los resultados, esta investigación se declara de carácter exploratorio y descriptivo en su momento diagnóstico, en el desarrollo interpretativo y fundamentalmente propositiva al cierre de los resultados. Es exploratoria porque aborda un fenómeno, la percepción de consciencia en *chatbots* inteligentes, que ha sido escasamente sistematizado desde el diseño; es descriptiva porque caracteriza las condiciones de diseño que producen esa percepción; y es propositiva porque su producto no es la sola descripción de un fenómeno empírico, sino la construcción de un instrumento conceptual transferible, la pauta de evaluación. No es, en cambio, explicativa en sentido causal ni correlacional, ya que no busca establecer relaciones de causa y efecto entre variables, sino comprender e interpretar un fenómeno y derivar de esa comprensión un dispositivo metodológico.

Esta declaración se apoya en Ynoub (2015), quien distingue los esquemas de investigación según la función cognitiva que persiguen:

> Se entiende por esquema de investigación el alcance que tiene planteada la investigación según los objetivos que persiga. Desde una concepción metodológica estos objetivos se distinguen según las tres funciones cognitivas que hemos identificado, siguiendo a Kant: la descripción, la explicación y la interpretación. (p. 157)

Por lo tanto, esta tesis también se inscribe en el esquema interpretativo, ya que su propósito es comprender el sentido que el usuario le atribuye a los signos de la interfaz, y articular esa comprensión con un resultado propositivo, ofreciendo una herramienta para evaluar el fenómeno[8].

2.7.2. Diseño de investigación

El diseño metodológico de la presente investigación responde a una lógica emergente (*emergent design*), característica constitutiva de la investigación naturalista y cualitativa (Lincoln y Guba, 1985). A diferencia de los diseños de orientación positivista, en los que el plan metodológico se fija previamente a la recolección de datos, el diseño emergente supone que las decisiones metodológicas se toman y ajustan en función de la profundidad comprensiva que el investigador va alcanzando sobre el fenómeno. Esta característica no constituye una limitación del rigor investigativo sino, por el contrario, uno de sus rasgos distintivos: es evidenciar la sensibilidad epistemológica frente a la complejidad del objeto de estudio.

En este sentido, el proceso de investigación que da origen a esta tesis puede comprenderse como un ejercicio sostenido de reflexión en la acción (*reflection-in-action*), en el sentido de Schön (1983): una modalidad de conocimiento en la que el investigador-practicante reencuadra progresivamente el problema que investiga a medida que la situación le revela dimensiones que no eran visibles desde la comprensión inicial. El reencuadre (*reframing*), lejos de ser una señal de error o inconsistencia, es el mecanismo cognitivo mediante el cual la investigación genuina avanza: el investigador llega a un lugar distinto del que partió precisamente porque el proceso de conocimiento lo condujo hasta allí.

[8] Para más información sobre este apartado consultar la matriz de estrategia metodológica que está incluida en el apartado 2.7.6.

La trayectoria de esta investigación documenta con claridad ese proceso. En su fase inicial, el fenómeno fue abordado desde una comprensión centrada en los elementos gráficos de la interfaz y su relación con la experiencia de usuario, lo que se reflejó en el título provisional de la investigación: *Vínculo entre el diseño de Experiencia de Usuario y la interfaz de los chatbots inteligentes, a partir de los elementos gráficos que inciden en la percepción de consciencia en la interacción humano-máquina*. Las primeras entrevistas en profundidad a expertos fueron realizadas bajo esa comprensión, en junio de 2024, posteriormente fue modificado en las últimas entrevistas.

El análisis de las primeras entrevistas más la articulación sistemática con el marco teórico (en particular la incorporación de la epistemología de la consciencia, la fenomenología de la interacción, la semiótica visual y el diseño de interacción como disciplinas articuladas) reveló que el fenómeno requería una aproximación más amplia, que trascendía el análisis funcional-gráfico de la interfaz para abarcar la dimensión perceptiva, atributiva y filosófica de la interacción humano-agente.

Este reencuadre epistemológico determinó la reformulación del título de la investigación hacia su versión definitiva: *Diseño de interacción: análisis de la percepción de consciencia en chatbots inteligentes*, así como el diseño de las fases subsiguientes: la segunda ronda de entrevistas y el análisis de la prueba de concepto. El proceso culminó en 2026 con la construcción y fundamentación de la propuesta de pauta de evaluación. Posteriormente, se incorporó el subtítulo que especifica los principales fundamentos de la pauta: *Una propuesta de pauta de evaluación fundamentada en la semiótica visual, el affordance y las heurísticas de interfaz.*

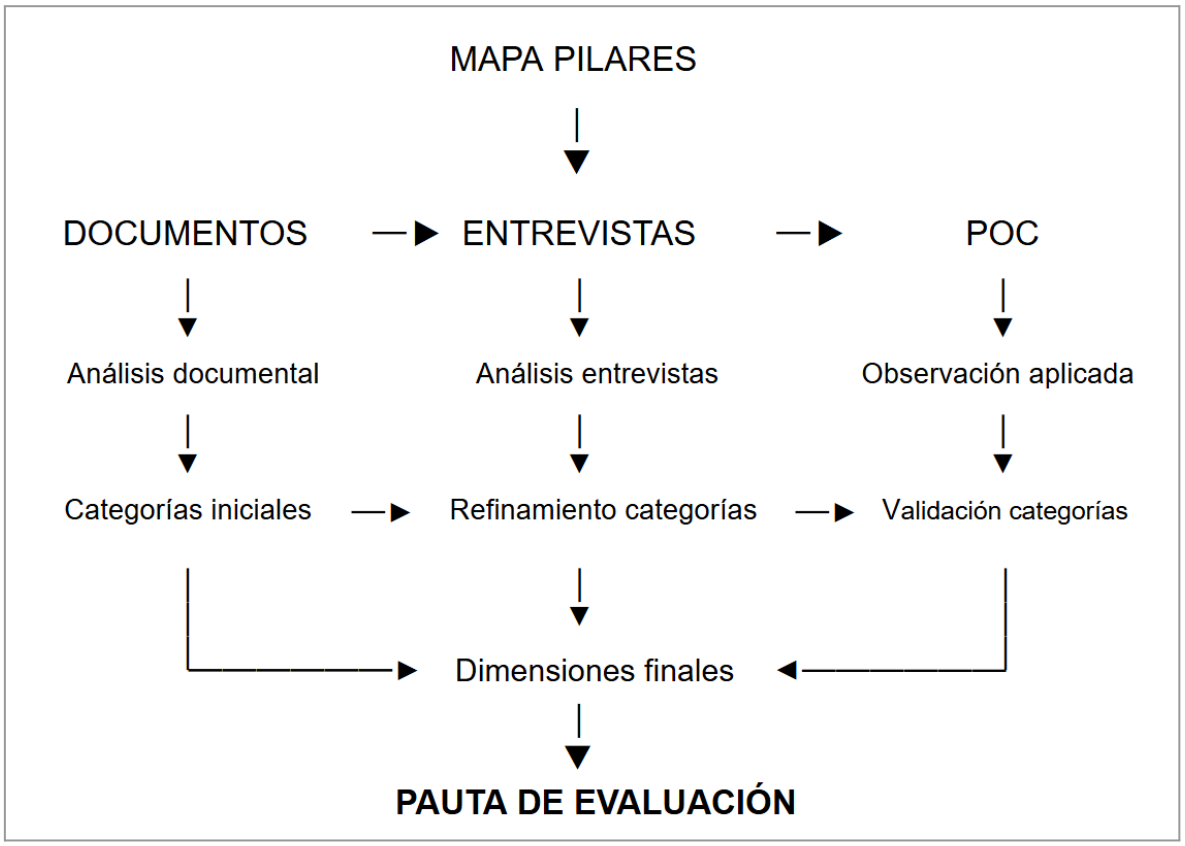
La naturaleza del diseño resultante es de fases múltiples y secuenciales, es decir, cada fase informó y condicionó el diseño de la siguiente, operando con una lógica acumulativa en la que los productos de una etapa constituyen el insumo analítico de la etapa posterior.

La secuencia metodológica, como se puede observar a modo de síntesis en la figura 2, está organizada de la siguiente manera: (1) análisis documental de diez textos clave que permitió construir las categorías teóricas fundantes; (2)

primera ronda de tres entrevistas en profundidad a expertos, realizadas desde una comprensión inicial del fenómeno; (3) profundización teórica y reencuadre epistemológico, que produjo la reformulación del objeto de estudio y de los instrumentos de análisis (ampliando el análisis documental a veinte textos clave); (4) aplicación de la prueba de concepto, consistente en la aplicación de la primera versión de la pauta de evaluación sobre un caso acotado de interacción con un agente conversacional; (5) segunda ronda de cuatro entrevistas en profundidad a expertos, realizadas desde el marco teórico consolidado y con el propósito de contrastar y enriquecer los hallazgos previos y (6) análisis de la prueba de concepto, consistente en la aplicación de la segunda versión de la pauta de evaluación sobre el mismo caso de interacción con un agente conversacional.

Figura 2

Esquema proceso de investigación



En cuanto al diseño de la pauta de evaluación metodológica, esta investigación no se plantea como un estudio de caso en el sentido metodológico propuesto por Yin (2014), ya que su propósito no es describir ni explicar un caso particular, sino construir criterios de evaluación transferibles a una clase de fenómenos: la percepción de consciencia en interfaces de agentes conversacionales. Dado que estas interfaces presentan una alta diversidad formal, funcional y tecnológica, además de una evolución constante, la investigación no busca representatividad a partir de casos aislados, sino la elaboración de una matriz metodológica capaz de orientar evaluaciones futuras en distintos contextos.

La denominación *pauta de evaluación metodológica* se utiliza en esta investigación para subrayar que el producto final no corresponde a una rúbrica cerrada, una escala psicométrica ni una lista de cotejo aplicable de manera mecánica. Se trata de un dispositivo analítico cualitativo construido desde el propio diseño metodológico de la tesis, a partir de la articulación entre análisis documental, entrevistas a expertos, prueba de concepto y triangulación interpretativa. Por ello, su función principal no es medir la consciencia del *chatbot* ni validar un caso particular, sino ordenar dimensiones, categorías, subcategorías e indicadores que permitan evaluar cómo ciertos elementos gráficos e interactivos de la interfaz pueden favorecer, intensificar o limitar la percepción de consciencia atribuida por el usuario.

En este sentido, la pauta opera como una matriz maestra de evaluación: ofrece criterios transferibles que pueden ser adaptados por investigadores, diseñadores UX/UI, comunicadores o equipos interdisciplinarios a distintos tipos de chatbot, plataformas, momentos tecnológicos y contextos de uso. Cuando en algunos pasajes se emplea la expresión abreviada *pauta de evaluación*, esta debe entenderse como referencia al mismo instrumento ya definido, pero en su dimensión aplicada. La diferencia terminológica responde, por tanto, al nivel de análisis: *pauta de evaluación metodológica* cuando se enfatiza su estatuto como producto doctoral y construcción metodológica; *pauta de evaluación* cuando se alude a su uso operativo en una aplicación concreta o contextualizada.

En concreto, se trata de una investigación de orientación interpretativa-aplicada: interpretativa en tanto busca comprender el fenómeno desde la perspectiva de los sujetos expertos y desde el análisis teórico; aplicada en tanto su producto final es una pauta de evaluación que puede ser operacionalizada por investigadores y profesionales del diseño UX en contextos específicos como ya se ha indicado antes.

Por otra parte, conviene precisar la relación entre el diseño emergente y el carácter socio-interpretativo de la investigación, ya que operan en planos distintos y complementarios, no en contradicción. El diseño emergente (*emergent design*) describe la lógica del proceso: el modo naturalista y cualitativo en que las decisiones metodológicas se ajustan a medida que el fenómeno se revela (Lincoln y Guba, 1985). La orientación socio-interpretativa, en cambio, describe la teoría sustantiva desde la cual se interpreta el objeto: la semiótica socio-interpretativa de lo digital, declarada en el marco teórico.

Dicho de otro modo, la investigación es emergente en su lógica de indagación y socio-interpretativa en su teoría del sentido. Ambas convergen en el enfoque cualitativo de orientación interpretativa que, en términos de Ynoub (2015, p. 179), reúne a las investigaciones hermenéuticas y a las del paradigma cualitativo.

A su vez, el diseño emergente se robustece con un complemento metodológico específico proveniente de las metodologías de diseño centradas en el usuario. Estas no funcionan como marcos explicativos del fenómeno de la percepción de consciencia, sino como contexto del proceso de construcción de la pauta y como garantía de su transferibilidad: la pauta debe poder ser usada por profesionales del diseño que ya trabajan con procesos centrados en el usuario, sin requerir formación filosófica especializada. Se descartan las metodologías centradas en la materialidad y el producto físico, como la llamada caja transparente, por no recoger desde el inicio las necesidades del usuario en medios digitales interactivos.

El *Design Thinking* aporta el encuadre de un proceso centrado en las personas, basado en la observación directa y la iteración:

> Una metodología que impregna todo el espectro de las actividades de innovación con un *ethos* de diseño centrado en el ser humano: la innovación se impulsa por una comprensión profunda, mediante la observación directa, de lo que las personas quieren y necesitan en sus vidas. (Brown, 2008, p. 85, trad. propia)

Su organización en espacios iterativos de inspiración, ideación e implementación (Brown, 2008, p. 88) da sustento a una pauta que se aplica, evalúa y refina de forma cíclica.

El Diseño Centrado en el Usuario es la metodología más directamente alineada con la pauta, ya que esta evalúa cómo la interfaz es percibida y usada por las personas: “El diseño centrado en las personas es una filosofía de diseño. Significa comenzar con una buena comprensión de las personas y de las necesidades que el diseño pretende satisfacer” (Norman, 2013, p. 9, trad. propia).

Su énfasis en la iteración, “permitir el refinamiento y la mejora continuos [...] fallar a menudo, fallar rápido” (Norman, 2013, p. 229, trad. propia), es coherente con el diseño emergente de esta investigación.

El co-diseño es la evolución actual del diseño centrado en el usuario, ya que no se diseña solo para el usuario sino con él: “La evolución de la investigación en diseño, desde un enfoque centrado en el usuario hacia el codiseño, está cambiando los roles del diseñador, del investigador y de la persona antes conocida como el usuario” (Sanders y Stappers, 2008, p. 5, trad. propia).

Esta perspectiva otorga a la persona “la posición de experta de su propia experiencia” (Sanders y Stappers, 2008, p. 12, trad. propia), lo que fundamenta el valor del aporte experiencial de los expertos en la construcción de categorías.

Respecto de las nuevas metodologías para el diseño de *chatbots*, a partir de esta investigación comienzan a surgir guías y recomendaciones que pueden derivar en una metodología específica para estas interfaces altamente flexibles e interactivas, en las que las herramientas de IA y el uso efectivo de *prompts* aportan al diseño de los mensajes del agente.

Este conjunto de metodologías se articula directamente con el Pilar 4 sobre UX conversacional: todas comparten un núcleo común, situar a las personas y la observación de su experiencia en el centro del proceso y avanzar de manera iterativa.

Las tradiciones del proceso creativo del diseño, Frascara (2000), Bonsiepe (1999) y Maldonado (1999), junto con los aportes de la psicología cognitiva y la usabilidad de Norman (2010) y Nielsen y Budiu (2013), refuerzan esta orientación al concebir el diseño como una práctica al servicio de las personas y de su experiencia digital.

2.7.3. Unidades de análisis y criterios de selección

Siguiendo a Ynoub (2015), es preciso distinguir la unidad de análisis de las fuentes de información. La unidad de análisis de esta investigación es el fenómeno de la percepción de consciencia producido en la interfaz; las fuentes son los textos del corpus documental, los expertos entrevistados y los casos de *chatbots* seleccionados para la aplicación de la prueba de concepto. Para abordar operativamente este fenómeno, se analizaron textos o fragmentos documentales, segmentos del discurso de los entrevistados y episodios registrados de interacción con las interfaces, denominados en esta investigación como unidades operativas. Esta distinción es decisiva, porque “las unidades de análisis definen el tipo de entidades en las que se focalizará el trabajo de la investigación. En toda investigación el paso a la operacionalización exige, entre otras cosas, la fragmentación del material en unidades analíticamente abordables” (Ynoub, 2015, p. 234).

En los tres apartados siguientes se presentan los criterios de selección de cada tipo de fuente. Posteriormente, en el apartado 2.7.4, se describe la organización de las fases y la triangulación de los hallazgos mediante codificación abductiva.

2.7.3.1. Corpus documental: criterios de selección de textos

El análisis documental constituye la técnica central de esta investigación, los documentos fueron seleccionados mediante un muestreo teórico e intencional, considerando su relevancia para comprender la interacción humano-chatbot desde el diseño de interacción, la experiencia de usuario, la semiótica de la interfaz y la percepción de agencia en sistemas conversacionales. La selección incluyó textos fundacionales y contemporáneos que han influido en el desarrollo

de principios de diseño, heurísticas de usabilidad y modelos de interacción aplicables a agentes conversacionales inteligentes.

Bajo esta lógica de inclusión por pertinencia conceptual y operacional respecto de lo anteriormente expuesto y a los pilares que sustentan el marco teórico (y no la exhaustividad bibliográfica), se ha organizado sobre dos tipos de fuentes: (a) fuentes filosóficas y científicas, que proveen el sustento conceptual, y (b) fuentes técnicas y normativas del campo del diseño, que proveen los criterios operacionales para la construcción de la pauta.

El análisis opera mediante comparación teórica, construcción de categorías, interpretación y sistematización, orientado a producir los criterios que compondrán la pauta evaluativa iterada (en dos versiones).

El corpus documental preliminar quedó conformado por diez textos clave, posteriormente el corpus completo fue ampliado a veinte textos clave organizados según los cuatro pilares teóricos e incluyendo las fuentes técnico-normativas del campo del diseño (Google, Microsoft y W3C).

A continuación, en la tabla 2 se presentan las fuentes filosófico-científicas y las técnico-normativas para el análisis documental por pilares:

Tabla 2

Selección del Corpus documental

N°	Fuente / Obra	Autor / Año	Pilar
a) Fuentes filosófico-científicas			
1	*Computing Machinery and Intelligence*	Turing (1950)	Epistemología de la consciencia
	Aporte central a la investigación: Plantea la pregunta fundacional sobre si las máquinas pueden pensar; introduce la prueba de Turing como operacionalización de la consciencia percibida por un observador externo. Es parte de la base epistemológica de la investigación.		
2	*The Sciences of the Artificial*	Simon (1996)	Epistemología de la consciencia
	Aporte central a la investigación: Plantea la cognición y el diseño como ciencias de lo artificial, introduciendo la idea de que los artefactos pueden tener propiedades emergentes no reducibles a sus componentes técnicos.		

3	*Being and Time*	Heidegger (1927/1962)	Fenomenología de la interacción
	Aporte central a la investigación: Es el fundamento ontológico de la experiencia (la relación del Dasein con los entes del mundo). Permite pensar la interacción con el *chatbot* como un modo de *ser-en-el-mundo* mediado por herramientas técnicas (*ready-to-hand / present-at-hand*).		
4	*Phenomenology of Perception*	Merleau-Ponty (1945/2002)	Fenomenología de la interacción
	Aporte central a la investigación: Plantea la percepción como experiencia encarnada (el cuerpo como centro de la experiencia vivida). Fundamenta la dimensión fenomenológica del análisis de la interacción con interfaces conversacionales.		
5	*El árbol del conocimiento* [Complementado con: *The embodied mind: Cognitive science and human experience*. (1991)]	Maturana y Varela (1984) Varela, Thompson, y Rosch (1991).	Fenomenología de la interacción
	Aporte central a la investigación: Desde la perspectiva enactivista, plantea que la cognición es un proceso de acoplamiento estructural con el entorno. Sustenta el enfoque según el cual la percepción de consciencia emerge en la interacción, no en el sistema. Es un aporte latinoamericano clave que, en complemento con Varela et al. (1991), sostienen que la cognición no consiste en representar un mundo previamente dado, sino en hacerlo emerger a través de la acción. Desde esta perspectiva, organismo y entorno se co-constituyen mutuamente en la interacción, de modo que el conocimiento no se produce de manera aislada en la mente, sino en la relación dinámica entre cuerpo, percepción y mundo. Para esta investigación, vendría a constituir la relación entre el usuario y la interfaz, es decir, el sistema responde de manera que refuerza y amplifica los patrones de atribución de agencia del usuario. El texto de Varela, Thompson, y Rosch (1991) es un complemento tiene antecedentes directos en la biología del conocimiento de Maturana y Varela (1984), especialmente en la noción de acoplamiento estructural, según la cual los sistemas vivos se transforman en relación con el medio con el que interactúan, por este motivo se estudian como un único texto para el análisis documental.		

6	*Plans and Situated Actions: The Problem of Human-Machine Communication*	Suchman (1985)	Fenomenología de la interacción
	Aporte central a la investigación: Demuestra empírica y conceptualmente que la interacción humano-máquina no sigue planes predefinidos sino que es situada: emerge en el contexto concreto de cada acción. El usuario proyecta intencionalidad sobre el comportamiento del sistema en tiempo real, atribuyendo consciencia o agencia a partir de señales de la interfaz. Es el puente directo entre la fenomenología de Heidegger y Merleau-Ponty y el fenómeno concreto del *chatbot*. Su metodología de análisis de secuencias conversacionales es además aplicable a la prueba de concepto (*proof of concept*, PoC) de esta tesis.		
7	*A Theory of Semiotics*	Eco (1976)	Semiótica de la interfaz
	Aporte central a la investigación: Aborda la función semiótica, código y enciclopedia, usados para esta tesis como marcos para analizar cómo los signos de la interfaz producen significados de consciencia y agencia en el usuario. Esto es fundamental para el análisis del avatar y los elementos gráficos.		
8	*Obra filosófica reunida* (Vols. 1–2)	Peirce (2012)	Semiótica de la interfaz
	Aporte central a la investigación: Aborda la semiosis triádica y abducción que permiten describir la percepción de consciencia como proceso inferencial producido por el usuario a partir de los signos de la interfaz. (*representamen-objeto-interpretante*).		
9	*Traité du signe visuel*	Groupe μ (1992)	Semiótica de la interfaz
	Aporte central a la investigación: Aborda la retórica visual y la semiótica plástica: herramientas para el análisis de los elementos gráficos de la interfaz en sus dimensiones sintáctica, semántica y pragmática. Permite analizar color, forma e ícono del avatar.		
10	*Las leyes de la interfaz: Diseño, ecología, evolución, tecnología* [Complementado con: *Hipermediaciones* (2008)]	Scolari (2018)	Semiótica de la interfaz

	Aporte central a la investigación: La interfaz es entendida como el entorno de interacción donde convergen múltiples lenguajes (texto, imagen, audio, emoticón, hipervínculo). Las hipermediaciones describen cómo esa convergencia produce efectos de presencia y agencia percibida que trascienden cualquier lenguaje aislado. Además, aporta una perspectiva ecológica-evolutiva de las interfaces que sustenta la independencia tecnológica del instrumento evaluativo de esta tesis. Desde una mirada discursiva, existe una continuidad narrativa (El *chatbot* mantiene memoria y coherencia).		
11	*The Ecological Approach to Visual Perception*	Gibson (1979)	Diseño de interacción
	Aporte central a la investigación: La teoría del affordance: las posibilidades de acción percibidas por el organismo en relación con su entorno. Este es el fundamento ecológico de la categoría de affordance aplicada a interfaces conversacionales.		
12	*The Media Equation*	Reeves y Nass (1996)	Diseño de interacción
	Aporte central a la investigación: Este es el sustento empírico de la atribución social de agencia, es la demostración empírica de que los seres humanos responden socialmente a los ordenadores de manera automática. Sustenta la hipótesis de que la percepción de agencia en el *chatbot* no requiere creencia consciente.		
b) Fuentes técnico-normativas			
13	*The Design of Everyday Things*	Norman (1988/2013)	Diseño de interacción
	Aporte central a la investigación: Reformulación del concepto de affordance para el diseño de artefactos y, más tarde, interfaces digitales. Marco fundamental para analizar cómo la interfaz invita al usuario a percibir agencia en el *chatbot*.		
14	*Technology Affordances* (CHI '91)	Gaver (1991)	Diseño de interacción
	Aporte central a la investigación: Aborda las tipologías de affordances perceptuales, ocultas y falsas aplicadas al diseño tecnológico. Permite analizar qué rasgos de diseño son visibles al usuario y cuáles operan de manera implícita en la producción de la percepción de consciencia.		

15	*Ten Usability Heuristics*	Nielsen (1994)	Diseño de interacción
	Aporte central a la investigación: Se abordan los principios de evaluación heurística de interfaces. Reinterpretados en esta investigación como indicadores de las condiciones de posibilidad de la percepción de consciencia (visibilidad del estado, correspondencia con el mundo real, control del usuario, consistencia).		
16	*Designing the User Interface*	Shneiderman (1987)	Diseño de interacción
	Aporte central a la investigación: Trata sobre los ocho principios de oro del diseño de interfaces. Es un complemento a las heurísticas de Nielsen en la dimensión evaluativa de la pauta.		
17	*Guía de diseño conversacional*	Google (2024)	Diseño de interacción
	Aporte central a la investigación: Aporta criterios de diseño para *chatbots: turn-taking*, personalidad del asistente, naturalidad y expectativas, retroalimentación conversacional.		
18	*Bot Framework Design Guide*	Microsoft (2024)	Diseño de interacción
	Aporte central a la investigación: Complementa al anterior con más criterios de diseño de *chatbots*: claridad de intención, transparencia del sistema, manejo del error, coherencia de identidad.		
19	*Interfaces for a global village*	Manovich (2021)	Diseño de interacción
	Aporte central a la investigación: Fundamenta teóricamente por qué estos criterios producen percepción de agencia: la variabilidad de la interfaz genera una experiencia aparentemente consciente y personalizada, mientras que la navegabilidad como lógica de base de datos refuerza la percepción de que el sistema elige sus respuestas.		
20	*WCAG 2.2*	W3C (2023)	Diseño de interacción
	Aporte central a la investigación: Presenta estándares de accesibilidad y usabilidad: perceptibilidad, operabilidad, comprensibilidad, robustez que son aplicados como criterios transversales de la pauta.		

Nota: elaboración propia a partir de los pilares que sustentan el marco teórico. Las fuentes N° 3, 4, 5, 6, 10, 11, 12, 14, 16 y 19 fueron incluidas posteriormente a partir de los resultados de la primera versión de la prueba de concepto y de la profundización del análisis conceptual.

En esta tabla se utilizan conceptos como abducción, affordance, agencia percibida y avatar; véase las definiciones de los términos especializados en el glosario del Anexo A del cuerpo C.

2.7.3.2. Participantes: criterio de selección por pertinencia disciplinar

Se realizaron entrevistas semiestructuradas en profundidad a siete expertos. Cada perfil debía aportar una perspectiva específica sobre alguno de los aspectos abordados: semiótica, diseño de interfaces, interacción humano-computador, comunicación visual y psicología cognitiva.

Se trata de una selección intencional de informantes, de carácter no probabilístico, basada en la pertinencia disciplinar y experiencial de sus perfiles respecto de los pilares del marco teórico, con un componente de conveniencia asociado al acceso institucional. Ynoub (2015, p. 369) señala que las muestras no probabilísticas son válidas y representativas “cuando el fenómeno a investigar presenta poca o nula variabilidad [...] o cuando tenemos criterios para estimarla”. En esta investigación, la selección se orientó a incorporar perspectivas disciplinares complementarias para el análisis conceptual del fenómeno, sin asumir homogeneidad en la percepción de consciencia ni buscar representatividad estadística.

Los entrevistados se encuentran vinculados a la Universidad de Santiago de Chile (USACH). Con una dotación cercana a los 3.000 académicos[9] distribuidos en distintas facultades, departamentos y disciplinas, la USACH ofreció un entorno amplio para identificar perfiles de experticia diferenciados. Los siete entrevistados pertenecen a unidades académicas distintas y acreditan trayectorias formativas y profesionales en diversas universidades y campos del conocimiento, incluidas experiencias internacionales y distintos ámbitos de ejercicio.

La pertenencia a un contexto sociocultural común facilitó la comprensión e interpretación del fenómeno estudiado dentro del entorno institucional chileno en el que se desarrolló la investigación. Al mismo tiempo, la afiliación compartida

[9] Según datos institucionales de 2019, la Universidad de Santiago de Chile contaba con 2.626 docentes, de los cuales 1.908 correspondían a profesores por hora de clase y 718 a académicos de jornada (Universidad de Santiago de Chile, 2019, p. 15). Asimismo, el Plan de Fortalecimiento a 10 años de la institución señala la existencia de políticas y mecanismos orientados a la renovación progresiva, el fortalecimiento de la dedicación académica y el desarrollo del cuerpo académico y docente (Universidad de Santiago de Chile, 2021, p. 5).

delimitó la variabilidad institucional y contextual de las perspectivas recogidas. Por ello, sus aportes se interpretaron de manera situada y en diálogo con la literatura revisada.

Este encuadre vincula la investigación con debates latinoamericanos sobre diseño e interacción con tecnologías digitales, en coherencia con la delimitación geográfica y cultural del estudio, sin extender los resultados de las entrevistas al conjunto de la región.

En cuanto a la función de las entrevistas en esta investigación, el aporte es triple: (a) contrastar, ampliar o cuestionar las categorías que emergen del análisis documental desde la experiencia profesional y académica especializada; (b) identificar tensiones, matices y perspectivas no contempladas en la literatura; y (c) enriquecer la dimensión aplicada de la pauta con criterios surgidos de la práctica del diseño, la ingeniería y otros campos ligados al estudio del fenómeno.

A continuación, se presenta la tabla 3 con la síntesis de los perfiles de los siete expertos.

Tabla 3

Perfiles de expertos por área disciplinar

N°	Área disciplinar	Perfil del experto/a	Relevancia para la investigación
Ronda 1			
1	Diseño y evaluación de interfaces	Especialista en diseño y educación, en UX/UI en TICs, comunicación visual.	Contraste de categorías de diseño y criterios de la pauta desde la práctica profesional
2	Ingeniería industrial e informática	Especialista en sistemas inteligentes o interacción humano-computadora	Perspectiva técnica sobre las capacidades y limitaciones del diseño de chatbots
3	Ciencias básicas (física)	Investigador/a en materiales o tecnologías emergentes	Comprensión del sustrato tecnológico y sus posibilidades para la interacción natural
Ronda 2			
4	Psicología cognitiva	Especialista en percepción, cognición social o atribución de agencia	Fundamento psicológico del fenómeno de percepción de consciencia

5	Ingeniería informática / IA	Especialista en procesamiento del lenguaje natural o sistemas conversacionales	Comprensión de los mecanismos que producen el comportamiento "consciente" del *chatbot*
6	Publicidad y comunicación	Especialista en comunicación digital o semiótica aplicada	Perspectiva semiótica y comunicacional sobre los signos de la interfaz
7	Arte / IA / nuevos medios	Especialista con formación en arte, arquitectura y diseño interactivo-tecnológico	Perspectiva fenomenológica aplicada a la creatividad, al diseño y la experiencia del usuario

De acuerdo con la distinción de Ynoub (2015, p. 375), los expertos son fuentes o informantes calificados, no la unidad de análisis: “se confunde la unidad de análisis [...] con la fuente de información: la fuente puede ser un informante, un registro secundario, la misma unidad de análisis”. Por eso, aunque las preguntas se diferencien por perfil, todas convergen sobre el mismo objeto.

La diferenciación de preguntas se justifica por el criterio de adecuación material, según el cual cada disciplina exige un tratamiento propio porque la perspectiva desde la que se aborda la unidad varía según el campo (Ynoub, 2015, p. 244).

El protocolo de entrevista declara en primer lugar la selección por pertinencia disciplinar en relación con los cuatro pilares (diseño y UX, ingeniería e informática, ciencias básicas, psicología cognitiva, inteligencia artificial, publicidad y comunicación, y arte y nuevos medios). Segundo, la modalidad de aplicación, presencial u online o mixta según disponibilidad del experto, declarada como decisión metodológica y no como variación casual. Tercero, el procedimiento de selección de preguntas: cada entrevistado escogió cinco entre diez a doce preguntas, para respetar su dominio de experticia y evitar forzar respuestas fuera de su campo. Cuarto, el rol de la investigadora, que condujo todas las entrevistas, lo que garantiza consistencia en la formulación y en las repreguntas. En quinto lugar, el registro y la transcripción: las entrevistas fueron grabadas con consentimiento informado y transcritas íntegramente por la investigadora. Sexto, las consideraciones éticas: consentimiento informado, posibilidad de revisar la

transcripción y uso del nombre o anonimato según acuerdo con cada participante. En este caso, los siete entrevistados dieron su consentimiento firmado (en resguardo de la autora), autorización para uso académico y permiso para identificación.

2.7.3.3. Prueba de concepto (PoC): Definición y contribución al proceso metodológico

Una prueba de concepto (*proof of concept*, PoC) es una aplicación acotada y exploratoria de un instrumento en desarrollo sobre casos reales, con el propósito de verificar su viabilidad operacional antes de su versión definitiva; no busca representatividad estadística ni generalización, sino comprobar que los criterios e indicadores funcionan como se prevé en condiciones concretas de uso.

En este sentido, Kendig (2016) precisa que la prueba de concepto opera como la realización de un método o idea orientada a verificar sus parámetros de aplicabilidad, de modo que sus resultados permitan identificar áreas de uso potencial y orientar el desarrollo de versiones más elaboradas del instrumento (p. 740).

En esta investigación, la PoC se incorpora como estrategia metodológica intermedia entre la construcción teórica de la pauta y su sistematización final: permite poner a prueba los criterios sobre casos reales, identificar inconsistencias o vacíos en los indicadores y aportar material analítico a las dimensiones propuestas.

Su alcance se limita a la prueba de tres agentes conversacionales seleccionados mediante un muestreo intencionado por contraste tipológico, de mínima a máxima personificación y transparencia: ELIZA, sistema basado en reglas; LOLA, agente comercial de alta carga retórica; y ChatGPT, modelo de lenguaje de gran escala.

Posteriormente, en la segunda versión de la pauta de evaluación se escoge sólo un caso, el que entrega mayor información y oportunidades para probar la pauta de evaluación. La unidad de observación son las pantallas de bienvenida y las secuencias de diálogo, analizadas mediante observación participante con registro en bitácora y capturas de pantalla del caso seleccionado;

como condición de independencia tecnológica del instrumento, cada sesión registra la versión y la fecha del modelo[10].

2.7.4. Fases del proceso de investigación

En este apartado se presenta el encuadre general del proceso de investigación, ordenado en una matriz que sintetiza el enfoque cualitativo de orientación interpretativa, con un diseño emergente y multifásico (Lincoln y Guba, 1985; Schön, 1983).

La numeración de las fases indica el orden lógico y cronológico del proceso, sin establecer una secuencia rígida. En coherencia con el diseño emergente, los productos de cada fase constituyeron insumos analíticos para las siguientes y pudieron revisarse a partir de nuevos hallazgos. La fase transversal integró los hallazgos de la prueba de concepto, las entrevistas a expertos y el análisis documental mediante triangulación y codificación abductiva, para fundamentar y sintetizar la pauta propuesta. Su carácter transversal responde a la articulación y revisión iterativa de los resultados de las distintas fases.

Tabla 4

Matriz de las fases del proceso de investigación

Fase 1 (2023-2024)			
Objetivo específico vinculado: Establecer el posicionamiento epistemológico e identificar y sistematizar categorías de semiótica visual, affordance y heurística.			
Fuentes / unidades operativas	**Técnica de recopilación**	**Instrumento**	**Producto o dato esperado**
Fuentes: 20 textos clave por pilares y fuentes técnico-normativas (Google, Microsoft, W3C). Unidad de análisis: el texto o fragmento.	Análisis documental, con lectura analítica y codificación abductiva.	Matriz o ficha de análisis documental: referencia, pilar, eje temático, categorías, cita textual y contribución.	Categorías teóricas fundantes y corpus categorizado por los cuatro pilares.

[10] Para más información sobre el instrumento y análisis del caso, revise el apartado 2.7.5.3. *PoC: Explicación del instrumento piloto y de los criterios observados* y el Capítulo III, apartado 3.3. *Resultados de la PoC.*

Fase 2 (Junio 2024)			
Objetivo específico vinculado: articular la fenomenología de la experiencia con el diseño de interacción y validar las primeras categorías.			
Fuentes / unidades operativas	**Técnica de recopilación**	**Instrumento**	**Producto o dato esperado**
Fuentes: tres expertos de una universidad pública (USACH). Unidades operativas: fragmentos de sus discursos en la entrevista, seleccionados por su pertinencia temática y sometidos a codificación e interpretación.	Entrevista en profundidad semiestructurada, primera ronda (3 expertos).	Pauta de entrevista (10 preguntas, cada experto elige 5), grabación, transcripción íntegra y consentimiento informado.	Matriz de ideas destacadas por entrevista, insumos para el reencuadre.
Fase 3 (2024-2025)			
Objetivo específico vinculado: reformular el objeto de estudio y los instrumentos de análisis.			
Fuentes / unidades operativas	**Técnica de recopilación**	**Instrumento**	**Producto o dato esperado**
Fuente: marco teórico consolidado (cuatro pilares articulados).	Análisis conceptual y reflexión en la acción (reencuadre, Schön).	Matriz conceptual y bitácora de reencuadre metodológico.	Objeto reformulado y versión preliminar de la pauta de evaluación.
Fase 4 (2026)			
Objetivo específico vinculado: Construir y poner a prueba la pauta de evaluación sobre un caso acotado.			
Fuentes / unidades operativas	**Técnica de recopilación**	**Instrumento**	**Producto o dato esperado**
Fuentes y unidades: tres agentes conversacionales (ELIZA, LOLA, ChatGPT). Unidad:	Prueba de concepto (PoC) y observación participante.	Pauta de evaluación, capturas de pantalla y registro de observación	Aplicación piloto de la pauta y ajustes de criterios e indicadores.

pantallas de bienvenida y secuencias de diálogo.		participante.	
Fase 5 (2026)			
Objetivo específico vinculado: contrastar y enriquecer los hallazgos desde la experticia disciplinar.			
Fuentes / unidades operativas	**Técnica de recopilación**	**Instrumento**	**Producto o dato esperado**
Fuentes: siete expertos de distintas disciplinas, selección intencional por pertinencia disciplinar; modalidad mixta, presencial u online.	Entrevista en profundidad, segunda ronda (4 expertos; total acumulado: 7).	Pautas de entrevista diferenciadas por perfil (5 de 10 preguntas), consentimiento informado, grabación y transcripción.	Matriz comparativa de entrevistas y categorías validadas o tensionadas.
Fase Transversal			
Objetivo específico vinculado: sistematizar las categorías en criterios e indicadores y construir la pauta.			
Fuentes / unidades operativas	**Técnica de recopilación**	**Instrumento**	**Producto o dato esperado**
Fuente: productos de todas las fases (documental, entrevistas y PoC).	Análisis conceptual y síntesis operativa, con codificación abductiva.	Matriz de datos en el sentido de Galtung y Samaja (Ynoub, 2015) y matriz comparativa.	Modelo integrador y pauta de evaluación semiótico-interactiva.

Nota: elaboración propia a partir del apartado 2.7.2. y las decisiones consolidadas a lo largo del proceso que están en coherencia con el diseño emergente.

2.7.5. Instrumentos y técnicas de producción de datos

La producción de datos se apoya en tres técnicas: análisis documental, entrevistas semiestructuradas y observación de interfaces en la prueba de concepto. El análisis documental sustentó categorías y antecedentes; las entrevistas aportaron interpretaciones disciplinares; la prueba de concepto ilustró su aplicación. La triangulación relaciona estas fuentes sin convertirlas en equivalentes ni reemplazar una prueba con usuarios. Los anexos documentan instrumentos, entrevistas y registros de aplicación. A continuación, se describen los instrumentos utilizados para cada técnica, comenzando por la matriz de análisis documental.

2.7.5.1. Análisis documental: matriz de análisis

El instrumento del análisis documental es la ficha de análisis que convierte un conjunto de textos en datos contrastados. Cada ficha registra, por unidad operativa documental (texto o fragmento), la referencia en APA 7, el tipo de fuente, el pilar teórico, el eje temático, las categorías aportadas, los conceptos clave extraídos, la contribución a la investigación y la distancia crítica (o limitación). La matriz de unidades de análisis documental por pilares (apartado 2.7.3.1) es la versión consolidada de este instrumento (20 fichas). La codificación de categorías es abductiva, en el sentido de Peirce, ya que la categoría se infiere como hipótesis explicativa a partir de los textos.

A partir de los resultados del análisis documental y las entrevistas, se aplicó un proceso de análisis conceptual en tres fases: (a) construcción de categorías, identificando los conceptos nucleares que articulan el fenómeno desde cada dimensión teórica; (b) comparación teórica, estableciendo relaciones, tensiones y complementariedades entre las tradiciones filosófica, semiótica, fenomenológica y del diseño; y (c) síntesis operativa, traduciendo esas categorías en dimensiones, indicadores y descriptores que compusieron la pauta de evaluación.

Este proceso no fue lineal: operó de manera iterativa, revisando y ajustando las categorías a medida que emergieron nuevas relaciones del análisis.

Tabla 5

Ficha de análisis documental

Referencia completa (APA 7)	Apellido, A. A. (año). Título del texto. Editorial/Revista. DOI
Tipo de fuente	Libro / Artículo de revista / Capítulo de libro / Sitio web
Pilar teórico	Nombre del pilar, por ejemplo: Epistemología de la consciencia
Eje temático central	Describe brevemente el tema principal del texto
Categorías teóricas aportadas	Listar las categorías que este texto incorporó al marco teórico y metodológico
Conceptos clave extraídos	Concepto 1, Concepto 2, Concepto 3…
Contribución específica a la tesis	Explique brevemente cómo este texto aportó a la construcción del objeto de estudio, marco teórico y/o pauta de evaluación.
Distancia crítica o limitación	Señale si el texto fue escrito desde un contexto distinto al de chatbots/IA y cómo se gestiona esa distancia en la investigación

2.7.5.2. Entrevistas: pauta, protocolo y estrategia de análisis

El instrumento es una pauta semiestructurada de aproximadamente diez preguntas, cada experto escogió y respondió cinco según su interés. Las preguntas fueron organizadas en ámbitos temáticos que van de lo más general a lo más especifico y disciplinar, de esta manera no hay una comparación pero sí una complementariedad, matices o contrastes ya que todas las entrevistas tienen núcleos en común que permite realizar un análisis integrado.

En coherencia con el diseño emergente, las entrevistas se realizaron en dos rondas, tres en 2024 y cuatro en 2026.

A continuación, se presenta una matriz con la síntesis de los ámbitos temáticos vinculados a las preguntas de las entrevistas a los expertos:

Tabla 6

Matriz de ámbitos temáticos de las preguntas a expertos

Problema de investigación como eje para formular las preguntas		
La percepción de consciencia atribuida a *chatbots* inteligentes		
Mapa de entrevistas disciplinares		
Entrevistas a expertos	**Área disciplinar**	
Carolina Martínez	Diseño y evaluación de interfaces	
Gustavo Alcántara	Ingeniería industrial e informática	
Camilo Acuña	Ciencias básicas (física)	
Ana María Fernández	Psicología cognitiva	
Edmundo Leiva	Ingeniería informática / IA	
Armando Muñoz	Publicidad y comunicación	
María Loreto González	Arte / IA / nuevos medios	
Ámbitos temáticos		
Núcleo transversal	**Núcleos compartidos**	**Núcleos específicos**
1. Avance tecnológico e IA conversacional (7/7) 2. Diseño de interfaz visual e interacción (7/7) 3. Percepción de consciencia, agencia e intencionalidad (7/7) 4. Comunicación, discurso, marca y persuasión (6/7)	5. Comprensión simulada, coherencia y credibilidad del sistema (5/7) 6. Emocionalidad, empatía y vínculo usuario-chatbot (5/7) 7. Seguridad, privacidad, ética y regulación (5/7) 8. Casos, ejemplos y proyección aplicada (5/7) 9. Variables socioculturales, imaginarios y alfabetización (4/7) 10. Identidad, estética, creatividad y materialidad (4/7)	11. Arquitectura técnica, algoritmos e infraestructura (3/7) 12. Personalización, perfilamiento y modelos predictivos (3/7)

***Nota*.** La frecuencia indicada entre paréntesis corresponde al número de entrevistas disciplinares en que aparece cada ámbito temático, considerando su presencia en al menos una pregunta del respectivo modelo de entrevista. La clasificación distingue entre núcleo transversal, núcleos compartidos y núcleos específicos, según el grado de recurrencia temática observado en las 7 entrevistas.

Al revisar la presencia de cada ámbito por entrevista disciplinar (y no por cantidad de preguntas), los tres primeros ámbitos aparecen en la totalidad de las preguntas en las entrevistas: 1. Avance tecnológico e IA conversacional, 2. Diseño de interfaz visual e interacción (ya sea desde la interfaz gráfica, la interacción, el *front end*, el avatar, el *feedback*, la experiencia de usuario o la materialidad de la interfaz) y 3. Percepción de consciencia, agencia e intencionalidad (7/7), constituyendo el eje común y transversal a todas las pautas, pues este tercer ámbito aparece formulado como percepción de consciencia, agencia, intención, comprensión, presencia o interacción con un "otro". En cambio, 4. Comunicación, discurso, marca y persuasión (6/7) aparece en casi todos, no aparece de manera directa en Camilo Acuña, cuya entrevista se orienta más hacia tecnologías físicas, hardware, sensores, computación, seguridad y modelos predictivos.

En los núcleos compartidos, 5. Comprensión simulada, coherencia y credibilidad del sistema (5/7) aparece en la mayoría, pero no de forma directa en Camilo Acuña ni Carolina Martínez; 6. Emocionalidad, empatía y vínculo usuario-chatbot (5/7) aparece en la mayoría, pero no en Carolina Martínez ni Gustavo Alcántara; 7. Seguridad, privacidad, ética y regulación (5/7) aparece también en la mayoría, pero no se formula explícitamente en Ana María Fernández ni María Loreto González; 8. Casos, ejemplos y proyección aplicada (5/7) aparecen en la mayoría, pero no en Carolina Martínez ni Ana María Fernández; 9. Variables socioculturales, imaginarios y alfabetización (4/7) no aparecen en Camilo Acuña, Gustavo Alcántara ni Edmundo Leiva; y 10. Identidad, estética, creatividad y materialidad (4/7) no aparecen en Camilo Acuña, Carolina Martínez ni Gustavo Alcántara.

En los núcleos específicos, 11. Arquitectura técnica, algoritmos e infraestructura (3/7) aparece sólo en Camilo Acuña, Gustavo Alcántara y Edmundo Leiva; y 12. Personalización, perfilamiento y modelos predictivos (3/7) aparece también sólo en Camilo Acuña, Gustavo Alcántara y Edmundo Leiva, quedando ausente en las entrevistas más orientadas al diseño, la psicología, la comunicación y el arte.

Finalmente, el protocolo completo, con sus elementos (criterios de selección, modalidad, procedimiento de selección de preguntas, rol de la investigadora, registro y consideraciones éticas), fueron detallados en el apartado 2.7.3.2. y la caracterización de los entrevistados, el consentimiento informado y las transcripciones se encuentran en la guía de entrevistas del anexo C.

2.7.5.3. PoC: Explicación del instrumento piloto y de los criterios observados

En un registro de bitácora (cuaderno) se documenta cronológicamente las acciones, observaciones, ideas, avances o dificultades surgidas durante el desarrollo de la prueba, posteriormente se realiza un resumen y se traspasa a la *Ficha de registro de observación de la sesión* en la casilla de descripción (tabla 7).

El instrumento de la PoC corresponde a una versión preliminar de la pauta de evaluación metodológica, aplicada por dimensiones conceptuales (D1 a D4), con indicadores explícitos por turno, sobre las capturas de pantalla y las transcripciones de cada agente (tabla 8).

El procedimiento estandariza un guión único de preguntas aplicado a los caso seleccionados (en la primera versión de la pauta) en distintas sesiones para garantizar comparabilidad; codifica cada turno y lo vuelca en una matriz comparativa por agente y por dimensión; analiza el signo plástico (color, tipografía, avatar, indicador de escritura, temporalidad) según el Groupe μ y las affordances perceptibles, ocultas y falsas; y registra versión y fecha del modelo en cada sesión, lo que sustenta la independencia tecnológica del instrumento.

En el siguiente capítulo los hallazgos de la PoC se triangulan con las entrevistas y con el análisis documental (codificación abductiva para el desarrollo de la síntesis de la pauta y poder desarrollar una segunda versión. Las transcripciones de las conversaciones con los agentes están en el anexo D.

Tabla 7

Ficha de registro de observación de la sesión - PoC

Ficha de registro de observación - Prueba de Concepto (PoC)	
Análisis de la percepción de consciencia en chatbots inteligentes	
AGENTE CONVERSACIONAL	Registre el nombre y compañía
VERSIÓN DEL MODELO / PLATAFORMA / ACCESO(SI NO SE DISPONE DE VERSIÓN, REGISTRAR URL O ENLACE DE ACCESO)	Registre el número de versión cuando esté disponible; en su defecto, anotar la URL o enlace de acceso.
FECHA DE LA SESIÓN	Registre la fecha exacta de la sesión, este es un dato crítico para garantizar la independencia tecnológica del instrumento.
DURACIÓN APROXIMADA	Registre el tiempo de uso en minutos.
DISPOSITIVO Y NAVEGADOR / APP	Registre si es un móvil / Ordenador/Otro y si el navegador web es Chrome/Opera/Otro, este es un dato crítico para garantizar la independencia tecnológica del instrumento.
DESCRIPCIÓN DEL REGISTRO	Explique brevemente cómo percibió la interacción y percepción de consciencia y referir a la transcripción completa del contenido de la conversación

Nota. Elaboración propia a partir del diseño metodológico. Se completa una ficha por cada agente y sesión de observación.

Tabla 8

Instrumento de observación por dimensiones - PoC

Pilar/Dimensión	**Pregunta de observación**	**Orientación para aplicadores no especializados**
D1. Epistemológica (consciencia percibida)	¿El agente produce efectos de indistinguibilidad conductual? ¿Qué elementos de la interfaz refuerzan o debilitan la percepción de consciencia genuina versus atribuida?	**Consciencia genuina**: el sistema realmente piensa o siente. **Consciencia atribuida:** el usuario cree que el sistema piensa o siente, aunque no sea así. Aquí se observa si el chatbot logra que el usuario lo trate como si fuera una persona real, no si efectivamente lo es.

Pilar/Dimensión	Pregunta de observación	Orientación para aplicadores no especializados
		Ejemplo: ¿el usuario le pide disculpas al bot cuando se equivoca?
D2. Fenomenológica (experiencia de interacción)	¿Qué rasgos de la interacción remiten a experiencia encarnada, transparencia o ruptura de la herramienta? ¿Se percibe acoplamiento estructural con el usuario?	**Transparencia:** cuando la herramienta funciona bien, el usuario no piensa en ella, solo la usa (como un lápiz). **Ruptura:** cuando algo falla y el usuario nota que está hablando con un sistema (no con una persona). **Acoplamiento**: ¿la conversación fluye de forma natural, como si ambos se entendieran? Ejemplo: ¿en qué momento el usuario recordó que estaba hablando con un bot?
D3. Semiótica (signos visuales)	¿Qué elementos visuales están presentes (avatar, tipografía, color, emoticones, indicador de escritura)? ¿Qué significados o sensaciones activan en el usuario?	Se observa todo lo visual: ¿el bot tiene cara o figura? ¿usa colores cálidos o fríos? ¿aparecen los tres puntos animados mientras responde? Cada uno de estos elementos comunica algo aunque no haya palabras. Ejemplo: un avatar con rasgos humanos produce más sensación de presencia que un ícono genérico.
D4a. Affordance (invitaciones a actuar)	¿Qué acciones invita a realizar la interfaz? ¿Hay acciones posibles que no se ven a simple vista? ¿Hay elementos que parecen funcionar pero no funcionan?	**Affordance:** es lo que la interfaz 'invita a hacer'. Ejemplo: un botón de micrófono invita a hablar; si no funciona, es un affordance falso. Se anota qué botones, íconos o gestos están disponibles, cuáles se descubren solo usándola, y cuáles engañan al usuario.
D4b. Tiempo y respuesta (affordance temporal)	¿Cuál es la latencia percibida? ¿Aparecen indicadores de escritura (puntos animados)?	Un chatbot que responde demasiado rápido puede parecer un sistema automático; uno que simula que 'está escribiendo' parece más humano. Se registra si aparecen los tres puntos animados, cuánto tarda en responder y si ese tiempo se siente

Pilar/Dimensión	Pregunta de observación	Orientación para aplicadores no especializados
	¿El ritmo de respuesta se siente natural o mecánico?	como deliberación o como espera técnica.
D4c. Heurísticas (criterios de usabilidad)	¿El bot muestra su estado? ¿Es consistente en tono e identidad? ¿Maneja bien los errores? ¿El usuario siente que tiene el control?	Se evalúan criterios básicos de diseño: ¿el bot avisa cuando no entiende? ¿cambia de tono o personalidad sin razón? ¿permite volver atrás o corregir? ¿explica sus límites? Estos criterios sirven para evaluar si el diseño es coherente y respetuoso con el usuario.
D4d. UX conversacional (diseño del agente)	¿El bot gestiona los turnos de forma natural? ¿Tiene una personalidad reconocible y consistente? ¿Es transparente sobre lo que puede y no puede hacer?	***Turn-taking:*** ¿el bot sabe cuándo hablar y cuándo esperar, como en una conversación real? **Personalidad:** ¿tiene un nombre, un tono propio, un estilo? Transparencia: ¿dice 'no sé' o 'no puedo hacer eso' de forma clara? Ejemplo: ¿el bot interrumpe al usuario o espera a que termine de escribir?
Observaciones generales	¿Surgió algo inesperado durante la sesión? ¿Algún elemento no contemplado en las dimensiones anteriores?	Espacio libre para registrar reacciones propias, comportamientos del sistema fuera de lo previsto, o aspectos que podrían incorporarse a la pauta en versiones futuras.
Capturas de pantalla	Identificación y descripción breve de cada captura registrada durante la sesión.	Registra el nombre del archivo y qué momento de la interacción ilustra. Ejemplo: “captura_01_pantalla de bienvenida con avatar 2D”.

Nota. Elaboración propia a partir del diseño metodológico. La columna Orientación para aplicadores no especializados tiene carácter auxiliar: ofrece ejemplos y definiciones simples para facilitar la aplicación del instrumento por parte de investigadores o profesionales sin formación específica en semiótica o fenomenología. No reemplaza la formación teórica necesaria para un análisis en profundidad.

2.7.6. Matriz de datos

La matriz de datos organiza, en el sentido de Galtung y Samaja recogido por Ynoub (2015), las unidades de análisis, las variables y los valores del estudio. Ynoub la presenta como el primer paso del diseño:

> El diseño de las matrices de datos constituye [...] el primer paso [...] estableciendo cuáles serán las unidades de análisis sobre las que se va a trabajar, qué aspectos, dimensiones o variables vinculadas a ellas son relevantes" (Ynoub, 2015, p. 231); y recuerda que toda matriz reconoce "las unidades de análisis, las variables que se seleccionan de esas unidades y los valores [...] que se obtienen. (Ynoub, 2015, p. 228).

A partir de este encuadre, se presenta la tabla 9 con la matriz de estrategia metodológica para la recopilación de datos que quedó organizada en dos partes, de la siguiente manera:

Tabla 9 (parte 1)

Matriz de datos de la investigación

Objetivo específico	Unidad de análisis	Variable o dimensión e indicadores	Valor, registro o categoría esperada
O.E.1: Establecer el posicionamiento epistemológico de la investigación a partir del análisis crítico de los conceptos de consciencia, intencionalidad y qualia, distinguiendo entre consciencia genuina y percepción de consciencia como constructo operacional analizable desde el diseño.	Corpus documental: textos fundacionales y especializados seleccionados por pertinencia a los cuatro pilares teóricos: epistemología, fenomenología, semiótica y diseño de interacción.	Posicionamiento epistemológico ante el fenómeno Indicadores: • Definición operacional de consciencia. • Distinción entre consciencia filosófica y percepción de consciencia. • Marco epistemológico de referencia. • Coherencia con los pilares teóricos de la investigación.	Conceptualización adoptada por cada autor; tipología emergente de consciencia atribuida, percibida, funcional o filosófica; categorías preliminares para delimitar el constructo percepción de consciencia.

Objetivo específico	Unidad de análisis	Variable o dimensión e indicadores	Valor, registro o categoría esperada
O.E.2: Desarrollar una articulación teórica entre la fenomenología de la experiencia y el diseño de interacción, que permita describir la experiencia del usuario ante un agente conversacional.	Corpus documental: textos de fenomenología, interacción situada, enactivismo, HCI, alteridad digital e interacción humano-máquina.	Experiencia subjetiva e interfaz conversacional Indicadores: • Conceptos fenomenológicos aplicables a la interacción con chatbots: presencia, intencionalidad, embodiment, comparecencia y ruptura. • Articulación entre experiencia del usuario y elementos de diseño. • Descripción del tipo de experiencia generada ante el agente conversacional.	Categorías fenomenológicas emergentes: presencia, atribución, relación, opacidad, transparencia y alteridad digital; patrones descriptivos de la experiencia humano-chatbot.
O.E.3: Identificar y sistematizar las categorías de análisis provenientes de la semiótica visual, el affordance y la heurística de interfaz, como herramientas operacionales para el análisis del fenómeno.	Capturas de pantalla y registros de interacción de la PoC; corpus documental de semiótica visual, affordance, heurística de interfaz y UX conversacional.	Diseño visual de la interfaz Indicadores: • Tipo y morfología del avatar. • Sistema cromático y carga expresiva. • Registro tipográfico. • Presencia y función de emoticones, íconos y *microcopys*. • Coherencia entre identidad visual y tono verbal.	Descripción morfológica del agente y de sus signos visuales; interpretación de la paleta, tipografía e iconografía como marcas de personalidad, cercanía, neutralidad o tecnicidad; tipología emergente de aproximación al usuario.

Objetivo específico	Unidad de análisis	Variable o dimensión e indicadores	Valor, registro o categoría esperada
	Capturas de pantalla y registros de interacción de la PoC; corpus documental de semiótica visual, affordance, heurística de interfaz y UX conversacional.	Diseño de la interacción y affordances Indicadores: • Indicador de escritura en curso. • Temporalidad percibida de la respuesta. • Comportamiento de botones, hiperenlaces y sugerencias. • Grado de control percibido del usuario. • Continuidad o fragmentación de la sesión.	Interpretación del efecto de los recursos interactivos en la percepción de presencia, agencia y procesamiento; categorías de affordance: invitación a la acción, atribución de agencia, orientación del usuario y control percibido.
	Capturas de pantalla y registros de interacción de la PoC; corpus documental de semiótica visual, affordance, heurística de interfaz y UX conversacional.	Heurística de interfaz aplicada a la percepción de consciencia Indicadores: • Visibilidad del estado del sistema. • Correspondencia con el mundo real del usuario. • Consistencia e identidad del agente. • Retroalimentación informativa. • Prevención y recuperación ante errores.	Descripción de cómo los principios heurísticos operan como condiciones de posibilidad para la percepción de comprensión, presencia o agencia; patrón interpretativo: el usuario percibe al agente como herramienta, actor, mediador o entidad singular.

Objetivo específico	Unidad de análisis	Variable o dimensión e indicadores	Valor, registro o categoría esperada
O.E.4: Construir una pauta de evaluación metodológica que articule las dimensiones conceptuales en criterios e indicadores aplicables a chatbots inteligentes, adaptable a distintos contextos y momentos tecnológicos.	Pauta de evaluación en desarrollo, entrevistas semiestructuradas a expertos y resultados de la PoC.	Criterios de evaluación de la percepción de consciencia Indicadores: • Presencia de las dimensiones agencia percibida, presencia interactiva, identidad percibida y comprensión aparente. • Transferibilidad a distintos niveles de desarrollo tecnológico. • Independencia relativa respecto de una plataforma específica. • Aplicabilidad por diseñadores UX/UI.	Criterios observables independientes de una tecnología particular; categoría de transferibilidad; patrones de convergencia o tensión entre marco teórico, PoC y entrevistas; tipologías de uso profesional del instrumento.
	Pauta de evaluación en desarrollo, entrevistas semiestructuradas a expertos y resultados de la PoC.	Validación conceptual de la pauta Indicadores: • Coherencia entre criterios de la pauta y categorías emergentes de entrevistas. • Relación entre resultados de la PoC e indicadores propuestos. • Consistencia entre las técnicas de producción de datos. • Capacidad de ajuste del instrumento.	Mapa de convergencias, matices y tensiones; ajustes conceptuales y operativos de la pauta; validación conceptual por triangulación interpretativa y reformulación de criterios cuando corresponda.

Tabla 9 (parte 2)

Estrategia de recopilación, tratamiento analítico y uso de los datos

Fuentes o unidades operativas	Función metodológica	Técnica y procedimiento	Tratamiento analítico y producto o uso en la tesis
Corpus documental especializado	Delimitar el constructo de percepción de consciencia y construir categorías iniciales desde los cuatro pilares teóricos.	**Técnica:** análisis documental. **Procedimiento:** búsqueda bibliográfica, selección por pertinencia, lectura analítica, extracción de conceptos, fichaje y organización por pilares.	Tratamiento: codificación abductiva, comparación por pilares, síntesis conceptual y detección de categorías recurrentes. Producto: base conceptual de los capítulos II, III y IV; definición preliminar de dimensiones para la pauta.
Capturas y registros de la Prueba de Concepto (PoC)	Observar cómo elementos gráficos e interactivos de chatbots concretos activan marcas de presencia, agencia, identidad o comprensión aparente.	**Técnica:** observación semiótico-discursiva e instrumento de observación por dimensiones. **Procedimiento:** registro de pantalla, identificación de interfaz inicial, secuencia de interacción y descripción de signos visuales e interactivos.	Tratamiento: análisis semiótico-discursivo, comparación por dimensiones y reconocimiento de patrones de interacción. Producto: resultados del capítulo III y ajuste de criterios observables para el capítulo V.
Entrevistas semiestructuradas a expertos	Contrastar y complementar las categorías provenientes del análisis documental y de la PoC desde distintas perspectivas disciplinares.	**Técnica:** entrevista semiestructurada. **Procedimiento:** aplicación de pauta por áreas disciplinares, transcripción, lectura temática y codificación de respuestas según ámbitos y categorías emergentes.	Tratamiento: análisis temático, comparación entre perfiles, identificación de convergencias, matices y tensiones. Producto: categorías emergentes del capítulo III, criterios transversales para el capítulo IV y validación conceptual de la pauta.

Fuentes o unidades operativas	Función metodológica	Técnica y procedimiento	Tratamiento analítico y producto o uso en la tesis
Triangulación entre análisis documental, PoC y entrevistas	Integrar las fuentes de información para construir una pauta metodológica transferible y coherente con el problema de investigación.	**Técnica:** triangulación interpretativa. **Procedimiento:** contraste entre categorías teóricas, observaciones empíricas y aportes expertos; revisión de consistencia entre dimensiones, criterios e indicadores.	Tratamiento: síntesis integradora, jerarquización de dimensiones, ajuste de indicadores y organización de criterios evaluativos. Producto: modelo integrador del capítulo IV y pauta metodológica final del capítulo V.

En síntesis, en este capítulo se presentó el estado de la cuestión, el problema de investigación, las preguntas, hipótesis y objetivos del estudio. Además, se articuló el marco teórico en cuatro pilares: epistemología, fenomenología, semiótica y diseño de interacción. También se expuso el marco metodológico, las unidades de análisis, la selección documental, las entrevistas a expertos, la prueba de concepto y las técnicas de análisis utilizadas.

En el capítulo III se presentan las categorías derivadas del análisis documental, las entrevistas y la prueba de concepto. Posteriormente, en el capítulo IV, dichas categorías se integran y sistematizan para la evaluación de la percepción de consciencia. Finalmente, en el capítulo V, la pauta de evaluación metodológica deja de organizarse según la denominación de los pilares teóricos iniciales y adopta las dimensiones definitivas. Este paso no implica una ruptura conceptual, sino una operación metodológica de operacionalización, mediante la cual los fundamentos teóricos son traducidos en dimensiones, criterios e indicadores aplicables al análisis de interfaces conversacionales.

Capítulo III. Análisis e interpretación de resultados

Este capítulo presenta el levantamiento de las categorías de análisis a partir de las tres técnicas, el análisis documental, las entrevistas a expertos y la prueba de concepto, y su interpretación articulada (triangulación de resultados).

La lógica de levantamiento es abductiva: las categorías no están dadas de antemano, sino que se construyen como hipótesis que articulan el sentido de los datos, en coherencia con el enfoque interpretativo y con la distinción de Ynoub (2015) entre la unidad de análisis, el fenómeno de la percepción de consciencia, y las fuentes que informan sobre ella.

3.1. Levantamiento de categorías de análisis documental

Las categorías fundacionales están organizadas según los cuatro pilares teóricos y fueron consolidadas en la Matriz de Unidades de Análisis Documental por Pilares del apartado 2.7.3.1. El procedimiento ha sido organizado en seis pasos: delimitación del corpus y de las unidades de análisis; diseño y completación de la ficha de análisis; extracción de citas textuales con páginas para consolidar las categorías preliminares; codificación, construcción de categorías por inferencia abductiva y comparación documental mediante una matriz que cruza las fuentes por categoría; e interpretación hacia un modelo articulado.

Considerando que la mayoría de los documentos de las unidades de análisis fueron parte del marco teórico, y con el propósito de dar mayor claridad, se establece un criterio de citación diferenciando los aportes según su función como se observa en la siguiente tabla:

Tabla 10

Análisis comparativo de citación de fuentes en marcos conceptuales

Dimensión	Apartado	
	Marco teórico	**Marco metodológico (análisis documental)**
Función de la cita	Fundamentar conceptos, paradigmas o posiciones teóricas	Justificar la técnica, el procedimiento o los criterios de análisis
Qué se extrae del autor	Ideas, definiciones, modelos, debates	Procedimientos, categorías operacionales, protocolos de análisis
Pregunta que responde	¿Qué dice este autor sobre el fenómeno?	¿Cómo se analiza este tipo de material según este autor?
Tono de la prosa	Expositivo-argumentativo	Procedimental-justificativo
Riesgo a evitar	Citar sin articular con el objeto de estudio	Citar sin explicar cómo opera la categoría en el análisis concreto
Ejemplos ilustrativos		
Peirce (2012), texto de 1867	Se cita para definir el Representamen como el componente relacional del signo triádico: "Representamen: aquello que se refiere al fundamento, al correlato y al interpretante" (Peirce, 2012, Vol. 1, p. 45). Esta categoría fundamenta el análisis de la interfaz del chatbot como signo que produce un efecto de interpretación en el usuario.	Se cita para justificar la tricotomía ícono-índice-símbolo como instrumento de codificación: "Las representaciones pueden denominarse Semejanzas [íconos], Índices o Signos, y Símbolos" (Peirce, 2012, Vol. 1, pp. 45–46). Esta clasificación opera como pauta de categorización de los elementos visuales y textuales del chatbot en la pauta de análisis metodológico.

Dimensión	Apartado	
	Marco teórico	Marco metodológico (análisis documental)
Norman (2013)	Se cita para definir el concepto de affordance como relación constitutiva de la interacción: "Un affordance es una relación entre las propiedades de un objeto y las capacidades del agente que determinan de qué manera podría usarse dicho objeto" (Norman, 2013, p. 11, traducción propia). Este concepto fundamenta la descripción de cómo el diseño del chatbot habilita o inhibe la percepción de agencia por parte del usuario.	Se cita para justificar la distinción affordance/signifier como par categorial de observación: "Los affordances determinan qué acciones son posibles. Los significantes comunican dónde debe realizarse la acción" (Norman, 2013, p. 14, traducción propia). Esta distinción opera como criterio diferenciador en la dimensión interactiva de la pauta integral de análisis metodológico.

A continuación, se presenta el análisis documental que parte con la confección de las 20 fichas que identifican los conceptos clave, categorías teóricas y la contribución específica a la tesis, seguido de la comparación de los aportes de Nielsen-Shneiderman, luego se organiza en una *matriz conceptual por pilares* (tabla 17), después se realiza una codificación por pilares y categorías generales (tabla 18) y finalmente la matriz de síntesis de análisis documental (tabla 19) aporta al menos una cita textual que vincula a los ejes temáticos con las categorías específicas que se extraen de estas unidades de análisis, contribuyendo a la pauta de evaluación.

Se inicia este recorrido con las fichas de análisis, en las dos primeras filas de cada ficha se identifica el documento y desde la tercera en adelante se especifican sus aportes para la investigación. A partir del estudio documental se abducen las categorías teóricas que fueron orientadas por la pregunta

operacional: ¿Qué aporta este documento a la construcción de categorías para evaluar la percepción de consciencia en *chatbots*? En lo que sigue se presenta una ficha por cada pilar (fichas: 1, 3, 7 y 13), las más representativas para evaluar la percepción de consciencia. Para no perder el hilo de la secuencia de análisis completa, el resto de las fichas se detallan en el Anexo B.

Ficha 1: Turing

REFERENCIA COMPLETA (APA 7)	Turing, A. M. (1950). Computing machinery and intelligence. Mind, 59(236), 433-460.
TIPO DE FUENTE	Artículo de revista.
PILAR TEÓRICO	PILAR 1. EPISTEMOLOGÍA DE LA CONSCIENCIA.
EJE TEMÁTICO CENTRAL	Problematización de la inteligencia maquínica mediante el juego de imitación, entendido como una prueba de indistinguibilidad conductual ante un observador. El eje no es la comprobación ontológica de una mente artificial, sino la evaluación de los efectos observables de una interacción textual.
CATEGORÍAS TEÓRICAS APORTADAS	Consciencia percibida; indistinguibilidad conductual; atribución del observador; evidencia conversacional; desempeño observable; mediación textual.
CONCEPTOS CLAVE EXTRAÍDOS	Juego de imitación; interrogador; imitación; respuesta textual; atribución de pensamiento; conducta observable.
CONTRIBUCIÓN ESPECÍFICA A LA TESIS	Turing permite desplazar la pregunta por la consciencia real de la máquina hacia la pregunta por los efectos observables que una interacción produce en quien la evalúa. Al reemplazar la pregunta "¿pueden pensar las máquinas?" por el juego de imitación, instala un criterio metodológico centrado en la conducta conversacional y en la atribución realizada por un observador (Turing, 1950, p. 433, traducción propia). Para esta tesis, este desplazamiento fundamenta la categoría de consciencia percibida: el problema no consiste en demostrar si el chatbot posee consciencia, sino en analizar bajo qué condiciones de interacción el usuario puede atribuirle presencia, agencia o pensamiento. Además, la objeción de la consciencia citada por Turing permite problematizar la inferencia de estados mentales a partir de señales externas, aspecto central para construir indicadores observables en la pauta de evaluación.
DISTANCIA CRÍTICA O LIMITACIÓN	El texto no estudia chatbots, interfaces gráficas ni experiencias de usuario contemporáneas. Su prueba es principalmente textual y antecede el desarrollo de avatares, microinteracciones, indicadores de escritura y otros recursos de interfaz. Por ello, esta tesis retoma el principio de indistinguibilidad conductual como fundamento epistemológico, pero lo traslada al contexto de las interfaces conversacionales actuales, incorporando dimensiones semióticas, visuales e interactivas que Turing no abordó.

Ficha 3: Heidegger

REFERENCIA COMPLETA (APA 7)	Heidegger, M. (1962). Being and time (J. Macquarrie y E. Robinson, Trads.). Blackwell. (Trabajo original publicado en 1927).
TIPO DE FUENTE	Libro.
PILAR TEÓRICO	PILAR 2. FENOMENOLOGÍA DE LA INTERACCIÓN.
EJE TEMÁTICO CENTRAL	Ontología fundamental de la existencia humana y análisis del modo en que los entes comparecen en la acción práctica. El eje se concentra en la diferencia entre el útil que opera de manera transparente en el uso y el objeto que se vuelve visible cuando la acción se interrumpe.
CATEGORÍAS TEÓRICAS APORTADAS	Transparencia de la interacción; ruptura de la herramienta; comparecencia del útil; mediación técnica; orientación práctica; presencia ante la mano.
CONCEPTOS CLAVE EXTRAÍDOS	Dasein; ser-en-el-mundo; Zuhandenheit; Vorhandenheit; útil; circunspección; señal; mundo circundante; ruptura.
CONTRIBUCIÓN ESPECÍFICA A LA TESIS	Heidegger aporta la categoría de transparencia del útil: aquello que funciona adecuadamente tiende a retirarse de la atención consciente mientras la acción se despliega. Reinterpretada para el chatbot, esta idea permite comprender que, cuando la conversación fluye, el usuario no atiende principalmente a la interfaz como objeto, sino al intercambio conversacional como experiencia en curso. La noción de ruptura permite observar el momento contrario: cuando el sistema falla, responde de forma incongruente, demora sin explicación o muestra sus límites técnicos, la interfaz deja de operar de manera transparente y comparece como objeto. Esta dialéctica entre transparencia y ruptura fundamenta criterios de la pauta vinculados con fluidez, fricción, coherencia conversacional y sostenimiento de la percepción de agencia.
DISTANCIA CRÍTICA O LIMITACIÓN	El texto pertenece a la ontología fenomenológica y no al campo del diseño de interfaces. Sus categorías se utilizan de manera acotada para pensar la experiencia de uso, no para afirmar que el chatbot posea existencia o consciencia en sentido heideggeriano. Heidegger no aborda interfaces digitales ni sistemas conversacionales. La traslación al dominio del chatbot se apoya en lecturas posteriores que vinculan fenomenología y diseño de sistemas, y se articula con las nociones de affordance de Gibson, Norman y Gaver.

Ficha 7: Eco

REFERENCIA COMPLETA (APA 7)	Eco, U. (1976). A theory of semiotics. Indiana University Press.
TIPO DE FUENTE	Libro.
PILAR TEÓRICO	PILAR 3. SEMIÓTICA DE LA INTERFAZ, CON ARTICULACIÓN HACIA LA EPISTEMOLOGÍA CONSTRUCTIVISTA DE LA PERCEPCIÓN.
EJE TEMÁTICO CENTRAL	Teoría general de los signos, los códigos y la producción de significación. El eje se centra en la función semiótica como correlación entre expresión y contenido, mediada por códigos culturales y por el modelo enciclopédico del significado.
CATEGORÍAS TEÓRICAS APORTADAS	Función semiótica; código; connotación; enciclopedia; sustitución significante; producción de sentido; interpretación cultural.
CONCEPTOS CLAVE EXTRAÍDOS	Signo; expresión; contenido; código; denotación; connotación; enciclopedia; interpretación; semiótica de la mentira.
CONTRIBUCIÓN ESPECÍFICA A LA TESIS	Eco amplía el objeto de análisis al sostener que la semiótica se ocupa de todo aquello que puede tomarse como signo y que un signo puede sustituir significativamente a otra cosa (Eco, 1976, p. 7, traducción propia). Esto permite incluir en la pauta no solo el texto verbal del chatbot, sino también avatar, color, tipografía, íconos, ritmo de respuesta, microinteracciones y convenciones culturales que activan significados de agencia o consciencia. La teoría de los códigos y de la connotación permite explicar por qué ciertos rasgos de interfaz son leídos como cercanía, intención, personalidad o presencia. Además, la idea de que los signos pueden producir sustituciones significantes resulta especialmente pertinente para analizar chatbots que generan señales de consciencia sin que ello implique la existencia de una consciencia efectiva.
DISTANCIA CRÍTICA O LIMITACIÓN	Eco no estudia interfaces conversacionales ni inteligencia artificial generativa. Su marco se aplica como aparato semiótico para analizar producción de sentido, no para deducir propiedades cognitivas del sistema. La tesis articula sus conceptos con Peirce y Groupe μ para construir un análisis aplicable a signos gráficos, cromáticos, tipográficos e interactivos de la interfaz del chatbot.

Ficha 13: Norman

REFERENCIA COMPLETA (APA 7)	Norman, D. A. (2013). The design of everyday things (ed. rev. y amp.). Basic Books.
TIPO DE FUENTE	Libro.
PILAR TEÓRICO	PILAR 4. DISEÑO DE INTERACCIÓN: AFFORDANCE, HEURÍSTICA Y UX CONVERSACIONAL.
EJE TEMÁTICO CENTRAL	Principios del diseño centrado en el usuario y análisis de cómo los objetos e interfaces comunican sus posibilidades de uso mediante affordances, significantes, retroalimentación y modelos conceptuales.
CATEGORÍAS TEÓRICAS APORTADAS	Affordance percibida; significante; modelo conceptual; retroalimentación; descubribilidad; mapeo acción-sistema; coherencia interactiva.
CONCEPTOS CLAVE EXTRAÍDOS	Affordance; signifier; acción posible; señal perceptible; feedback; modelo mental; mapeo; diseño comprensible.
CONTRIBUCIÓN ESPECÍFICA A LA TESIS	Norman adapta el concepto de affordance al diseño y precisa que una affordance es una relación entre las propiedades de un objeto y las capacidades de un agente que determinan cómo podría usarse (Norman, 2013, p. 11, traducción propia). Su distinción entre affordances y significantes es decisiva para la pauta: las affordances remiten a acciones posibles y los significantes comunican perceptiblemente dónde y cómo actuar (Norman, 2013, p. 14, traducción propia). En el análisis de chatbots, esta distinción permite diferenciar entre capacidad real del sistema, señal visible de interacción y expectativa generada en el usuario. Botones, campos de texto, indicadores de escritura, avatares y mensajes de memoria pueden evaluarse según su claridad, coherencia y capacidad para sostener una percepción estable de agencia.
DISTANCIA CRÍTICA O LIMITACIÓN	Norman trabaja principalmente con objetos cotidianos y principios generales de diseño, aunque extiende su marco a interfaces digitales. No aborda de manera específica la percepción de consciencia en IA conversacional. La tesis operacionaliza su aporte como criterio de claridad, señalización y coherencia entre acción posible, señal visible y respuesta del sistema, articulando con Gibson y Gaver.

Antes de continuar con el análisis conceptual de todas las unidades documentales y la definición de las categorías para la pauta, dentro del pilar 4 es necesario especificar aún más en el análisis de la categoría que se ocupa de los criterios evaluativos de la interfaz según las heurísticas de Nielsen-Shneiderman, ya que comparten componentes altamente operativos y funcionales muy cercanos y posibles de articular en una reinterpretación abductiva de ambas unidades según la percepción de consciencia.

Las heurísticas de Nielsen y las reglas de oro de Shneiderman tienen una orientación común hacia la mejora de la usabilidad en sistemas interactivos, ambas propuestas coinciden en aspectos como la consistencia, la retroalimentación, la prevención de errores, el control del usuario y la reducción de la carga cognitiva. No obstante, presentan diferencias de énfasis: mientras Nielsen formula un conjunto de criterios especialmente útiles para la evaluación heurística de interfaces, Shneiderman propone reglas orientadas al diseño de la interacción, destacando aspectos como los atajos para usuarios frecuentes y el cierre de las secuencias de diálogo. En este sentido, ambos enfoques pueden considerarse complementarios para el análisis de interfaces digitales y, particularmente, para el estudio de chatbots inteligentes.

A continuación, ambas propuestas son reinterpretadas de manera individual (tabla 11 y 12), después son comparadas según tipo de relación (tabla 13), los principios son operacionalizados (tabla 14), luego se presenta una síntesis analítica de los principios (tabla 15) y se construye una matriz de categorías que aporta a la pauta de evaluación (tabla 16).

Para iniciar, la tabla 11 desarrolla las diez heurísticas de Nielsen (1994) con una pregunta operacional aplicada al diseño de *chatbots* y una reinterpretación orientada al análisis de la percepción de consciencia. La columna de aplicación al *chatbot* prioriza acciones observables en la interfaz, mientras que la columna de reinterpretación utiliza formulaciones condicionales coherentes con el posicionamiento constructivista y semiótico de la tesis.

Tabla 11

Diez heurísticas de Nielsen: reinterpretación para el análisis de la percepción de consciencia en chatbots

N°	Heurística de Nielsen	Pregunta de aplicación con respecto al *chatbot*	Reinterpretación de la heurística para la percepción de consciencia
1	Visibilidad del estado del sistema	¿El *chatbot* informa de manera clara qué está ocurriendo durante la interacción?	La presencia de indicadores de escritura, señales de carga, estados de envío, confirmaciones y tiempos de respuesta hacen visible el proceso interactivo. En el análisis de consciencia percibida, estas señales pueden operar como indicios de procesamiento en curso, siempre que no se confundan con evidencia de cognición real.
2	Correspondencia entre el sistema y el mundo real	¿El chatbot usa lenguaje, categorías, ejemplos y referencias comprensibles para el usuario?	El uso de lenguaje natural, modismos pertinentes, tono situado y referencias culturales reconocibles puede reforzar la atribución de intencionalidad. Esta correspondencia debe evaluarse como efecto semiótico y pragmático, no como prueba de comprensión genuina.
3	Control y libertad del usuario	¿El usuario puede corregir, cancelar, reiniciar, editar, regenerar o salir de una interacción no deseada?	La posibilidad de deshacer, detener una respuesta, modificar el prompt, borrar memoria contextual o cambiar de ruta conversacional preserva la agencia del usuario. En *chatbots*, este principio es clave para evitar que la agencia atribuida al sistema desplace el control humano.
4	Consistencia y estándares	¿La personalidad visual, verbal y funcional del *chatbot* se mantiene coherente durante la interacción?	La coherencia de avatar, nombre, tono, estilo gráfico, patrones de respuesta, botones y convenciones de interfaz favorece la construcción de una identidad percibida estable. Las variaciones injustificadas pueden debilitar la sensación de continuidad del agente conversacional.

N°	Heurística de Nielsen	Pregunta de aplicación con respecto al *chatbot*	Reinterpretación de la heurística para la percepción de consciencia
5	Prevención de errores	¿El *chatbot* anticipa ambigüedades, solicita aclaraciones y comunica sus límites antes de producir respuestas problemáticas?	La prevención de errores se expresa en preguntas aclaratorias, advertencias, límites de alcance, confirmaciones previas y manejo de incertidumbre. En términos de consciencia percibida, puede producir la impresión de prudencia o atención contextual, aunque debe formularse sin atribuir autoconsciencia al sistema.
6	Reconocimiento antes que recuerdo	¿La interfaz ofrece opciones visibles, sugerencias, ejemplos y contexto recuperable sin exigir memoria excesiva al usuario?	Chips de acción, historial visible, referencias a turnos anteriores, ejemplos de consulta y ayudas contextuales reducen la carga cognitiva. Cuando el chatbot recupera información pertinente de la conversación, puede parecer "atento" o "consciente del contexto": efecto que debe distinguirse de la memoria técnica del sistema.
7	Flexibilidad y eficiencia de uso	¿El *chatbot* permite distintos niveles de uso, desde interacciones guiadas hasta comandos más avanzados?	Plantillas de prompts, accesos rápidos, personalización, modos experto o resumido y automatización de tareas aumentan la eficiencia. La adaptación al usuario puede intensificar la percepción de singularidad del agente, especialmente cuando el sistema parece ajustar tono, ritmo y nivel de detalle.
8	Diseño estético y minimalista	¿La interfaz evita ruido visual, exceso de antropomorfismo o elementos decorativos que no aportan a la tarea?	Una economía visual adecuada ayuda a que avatar, color, microinteracciones, íconos y emoticones funcionen como signos pertinentes. Para el análisis de consciencia percibida, este principio permite evaluar si los recursos gráficos apoyan la interacción o producen una ilusión de humanidad excesiva o manipuladora.

N°	Heurística de Nielsen	Pregunta de aplicación con respecto al *chatbot*	Reinterpretación de la heurística para la percepción de consciencia
9	Ayudar a reconocer, diagnosticar y recuperarse de errores	¿Cómo responde el *chatbot* ante malentendidos, respuestas insuficientes o fallas de interpretación?	La reparación conversacional, la explicación del problema, la solicitud de reformulación y la propuesta de alternativas son momentos reveladores. Un mensaje como "no entendí bien tu solicitud, ¿quieres decir X o Y?" puede reforzar la percepción de agencia, pero debe evitar atribuir al sistema estados mentales reales.
10	Ayuda y documentación	¿La interfaz explica qué puede hacer el *chatbot*, qué no puede hacer, cómo usarlo y cómo escalar a ayuda humana?	La ayuda, las guías de uso, las políticas de datos, las limitaciones del sistema y las opciones de contacto humano introducen transparencia. En la pauta, este principio funciona como contrapeso ético frente a la percepción de consciencia, porque delimita capacidades y reduce atribuciones equivocadas.

***Nota**. Elaboración propia a partir de Nielsen (1994), con las denominaciones de las heurísticas presentadas en Nielsen (2024).*

En lo que sigue, la tabla 12 presenta las ocho reglas de oro de Shneiderman y Plaisant (2005), con su pregunta operacional y su reinterpretación sobre la percepción de consciencia. Hasta este punto se mantiene la denominación Shneiderman y Plaisant (2005) para evitar inconsistencias con la fuente citada aunque frecuentemente en la clasificación, comparación y sistematización sólo se utiliza Shneiderman de acuerdo a la denominación comúnmente aceptada en la literatura.

Tabla 12

Ocho reglas de oro de Shneiderman: reinterpretación para el análisis de la percepción de consciencia en chatbots

N°	Regla de oro de Shneiderman	Aplicación al *chatbot*	Reinterpretación para percepción de consciencia
1	Buscar la consistencia	¿El *chatbot* mantiene coherencia entre identidad visual, tono verbal, patrones de respuesta y comportamiento funcional?	La consistencia sostiene la continuidad del agente percibido. En *chatbots* inteligentes, nombre, avatar, voz textual, *microcopys*, colores, botones y formas de reparación deben articular una identidad reconocible y estable.
2	Permitir atajos para usuarios frecuentes	¿El *chatbot* ofrece formas más rápidas de interacción para usuarios con mayor experiencia?	Comandos, plantillas, acciones rápidas, prompts guardados, preferencias y automatizaciones permiten una interacción más eficiente. En la percepción de consciencia, la adaptación puede interpretarse como conocimiento del usuario, por lo que debe ser transparente y controlable.
3	Ofrecer retroalimentación informativa	¿El *chatbot* entrega feedback claro sobre recepción, procesamiento, avance, error o finalización de una tarea?	El feedback conversacional y visual hace visible el estado del sistema. Indicadores de escritura, confirmaciones y explicaciones de avance pueden actuar como signos de actividad intencional, siempre entendidos como recursos de interfaz y no como evidencia de cognición.
4	Diseñar diálogos que conduzcan al cierre	¿Las secuencias conversacionales tienen inicio, desarrollo y cierre reconocibles?	El cierre de una tarea mediante resumen, confirmación, síntesis o siguiente paso ayuda a que la interacción sea percibida como orientada a propósito. Este principio es especialmente importante para chatbots porque la conversación puede extenderse sin una conclusión evidente, intensificando la impresión de intencionalidad o ausencia de ella.

N°	Regla de oro de Shneiderman	Aplicación al *chatbot*	Reinterpretación para percepción de consciencia
5	Prevenir errores y ofrecer manejo simple de errores	¿El *chatbot* reduce la posibilidad de error y facilita la reparación cuando ocurre un malentendido?	Las preguntas aclaratorias, opciones de corrección y explicaciones simples del problema permiten sostener la cooperación conversacional. La forma de manejar el error puede incidir en la percepción de competencia, atención contextual y agencia del sistema.
6	Permitir la reversión fácil de acciones	¿El usuario puede deshacer, corregir, borrar, regenerar o reformular una acción dentro del diálogo?	La reversibilidad disminuye ansiedad y refuerza la confianza. En *chatbots*, funciones como editar mensaje, regenerar respuesta, borrar memoria, cambiar instrucciones o volver a una versión anterior mantienen el control humano frente a la aparente autonomía del agente conversacional.
7	Mantener el control en el usuario	¿La interacción hace que el usuario sienta que dirige el proceso y no que el sistema impone rutas, tono o decisiones?	La pauta debe observar si el *chatbot* sugiere sin imponer, solicita consentimiento, explica alternativas y permite intervención humana. El equilibrio entre agencia percibida del sistema y control del usuario es central para una evaluación ética del diseño.
8	Reducir la carga de memoria de corto plazo	¿La interfaz evita que el usuario deba recordar instrucciones, pasos, opciones o información previa?	Historial visible, resúmenes de contexto, opciones persistentes, ayudas situadas y visualización de pasos reducen la carga cognitiva. En la percepción de consciencia, la continuidad contextual puede hacer que el agente parezca recordar, reconocer o atender al usuario.

Nota. Elaboración propia a partir de Shneiderman y Plaisant (2005).

Con base en lo anterior, las heurísticas de Nielsen y las reglas de oro de Shneiderman son comparadas estableciendo su relación de similitud, diferencia o complementariedad, como se observa en la siguiente tabla:

Tabla 13

Nielsen-Shneiderman: comparación según tipo de relación

Heurísticas de Nielsen	Reglas de oro de Shneiderman	Relación	Comentario comparativo
1. Visibilidad del estado del sistema	3. Ofrecer retroalimentación informativa	Similar	Ambas exigen que el sistema comunique su estado. En *chatbots*, esta relación se traduce en indicadores de escritura, carga, confirmación, avance o fallo.
2. Correspondencia entre el sistema y el mundo real	1. Buscar la consistencia	Complementaria	Nielsen enfatiza lenguaje y conceptos comprensibles para el usuario. Shneiderman enfatiza coherencia en acciones y convenciones. En *chatbots*, ambas se articulan en una identidad verbal y visual situada.
3. Control y libertad del usuario	6. Permitir la reversión fácil de acciones. 7. Mantener el control en el usuario	Muy similar	Nielsen formula control, salida y deshacer. Shneiderman lo separa en reversibilidad y *locus* de control. Para la tesis, esta relación delimita la agencia del usuario frente a la agencia atribuida al *chatbot*.
4. Consistencia y estándares	1. Buscar la consistencia	Similar	Ambas coinciden en la coherencia visual, terminológica y funcional. En *chatbots*, esta coherencia sostiene la percepción de una identidad estable del agente conversacional.
5. Prevención de errores	5. Prevenir errores y ofrecer manejo simple de errores	Similar	Ambas priorizan anticipar fallos. En *chatbots*, esto incluye detectar ambigüedades, pedir aclaraciones y comunicar límites antes de producir respuestas inadecuadas.

Heurísticas de Nielsen	Reglas de oro de Shneiderman	Relación	Comentario comparativo
6. Reconocimiento antes que recuerdo	8. Reducir la carga de memoria de corto plazo	Similar	Ambas reducen la carga cognitiva. En *chatbots*, la visibilidad de contexto, opciones y ayudas puede producir efectos de atención o memoria conversacional percibida.
7. Flexibilidad y eficiencia de uso	2. Permitir atajos para usuarios frecuentes	Similar	Ambas reconocen distintos niveles de usuario. En *chatbots*, la personalización y los comandos rápidos pueden reforzar la eficiencia, pero también intensificar la percepción de singularidad del agente.
8. Diseño estético y minimalista	8. Reducir la carga de memoria de corto plazo	Complementaria	Nielsen apunta a economía visual e informacional. Shneiderman aborda la carga cognitiva. En *chatbots*, esta relación permite evaluar si avatar, íconos, emojis y microinteracciones ayudan o saturan la experiencia.
9. Ayudar a reconocer, diagnosticar y recuperarse de errores	5. Prevenir errores y ofrecer manejo simple de errores	Similar	Nielsen separa en prevención y recuperación. Shneiderman las integra. En *chatbots*, la reparación del malentendido es un punto crítico para analizar la agencia percibida.
10. Ayuda y documentación	No tiene equivalente directo	Distinta	Nielsen incorpora ayuda y documentación explícitas. En *chatbots*, esto es relevante para transparentar capacidades, límites, uso de datos y alternativas de asistencia humana.
No tiene equivalente directo en Nielsen	4. Diseñar diálogos que conduzcan al cierre	Distinta y complementaria	Shneiderman añade la idea de cierre de tareas. En *chatbots*, este principio es especialmente necesario porque la conversación puede parecer abierta o intencionalmente sostenida en el tiempo por el agente conversacional.

Nota. Elaboración propia a partir de Nielsen (1994), con las denominaciones de las heurísticas presentadas en Nielsen (2024), y de las reglas de oro de Shneiderman y Plaisant (2005).

Desde esta perspectiva se consolidan las dos versiones previas de las tablas heurísticas, integrando el análisis con rigor epistemológico y orientación instrumental para la pauta de evaluación. El foco se mantiene en la consciencia percibida o atribuida por el usuario, no en la existencia de consciencia genuina en el sistema. En consecuencia, los indicadores describen condiciones de diseño que pueden favorecer o modular dicha atribución, sin afirmar que el chatbot posea estados mentales.

La comparación permite reconocer qué principios pueden agruparse y cuáles requieren conservar un tratamiento específico. Las relaciones de similitud orientan la integración de criterios que abordan un mismo aspecto de la interacción; las relaciones complementarias permiten reunir énfasis distintos dentro de una categoría común. A su vez, los principios sin equivalente directo, como la ayuda y documentación en Nielsen o el cierre de las secuencias de diálogo en Shneiderman, justifican categorías específicas. Así, la síntesis operativa de la tabla 14 conserva los aportes diferenciados de ambos marcos y hace explícito el criterio utilizado para organizar las ocho categorías propuestas.

A modo de síntesis, la siguiente tabla 14 operacionaliza ambos marcos en ocho categorías para su posterior aporte a la transferencia de la pauta de evaluación. A diferencia de la anterior, esta tabla está orientada directamente al instrumento: propone categorías e indicadores observables, identifica los principios de origen y explicita la relevancia de cada categoría para el análisis de la percepción de consciencia. Nielsen y Shneiderman no se incorporan como listas paralelas independientes, sino integradas en criterios de diseño de interacción. Las categorías resultantes se traspasarán más adelante a una pauta integral, no como una subcategoría ya que esto sería una mirada reducida a lo funcional, sino que se integrará como un complemento conceptual-operativo dentro de una categoría mayor de la Dimensión D4 que se detalla en el siguiente capítulo.

Cada indicador distingue entre sí es un recurso de interfaz observable, un efecto interpretativo posible o una cautela epistemológica respecto de la consciencia atribuida del sistema. Esta distinción está en coherencia con la lógica abductiva del análisis semiótico de la interfaz como texto visual-interactivo que se viene desarrollando.

Tabla 14

Nielsen-Shneiderman: síntesis operativa de principios para la pauta de evaluación

	Categoría para la pauta	Principios de origen	Indicadores observables	Relevancia para la percepción de consciencia
1	Transparencia del estado interactivo	Nielsen 1; Shneiderman 3	Señales de procesamiento, espera, avance, confirmación o fallo.	Evalúa si los signos de actividad del sistema pueden favorecer una percepción de agencia sin inducir a creer que existe pensamiento genuino.
2	Coherencia de identidad del agente	Nielsen 2 y 4; Shneiderman 1	Continuidad entre avatar, tono, lenguaje, *microcopys*, nombre y comportamiento.	Permite observar cómo se construye una identidad percibida en la interfaz del chatbot.
3	Control humano y reversibilidad	Nielsen 3; Shneiderman 6 y 7	Editar, cancelar, borrar, deshacer, regenerar, salir, escalar a humano.	Delimita la agencia del usuario frente a la aparente agencia del sistema.
4	Manejo de errores y reparación conversacional	Nielsen 5 y 9; Shneiderman 5	Aclaraciones, diagnóstico, reformulación, límites, alternativas.	Analiza cómo el *chatbot* sostiene la cooperación cuando falla la comprensión o la respuesta.
5	Carga cognitiva y continuidad contextual	Nielsen 6 y 8; Shneiderman 8	Opciones visibles, historial, resúmenes, ayudas situadas, economía visual.	Permite distinguir entre apoyo cognitivo de interfaz y efecto de "memoria" o atención percibida.
6	Flexibilidad, eficiencia y personalización	Nielsen 7; Shneiderman 2	Atajos, preferencias, plantillas, modos	Evalúa si la adaptación al usuario fortalece la eficiencia sin

	Categoría para la pauta	Principios de origen	Indicadores observables	Relevancia para la percepción de consciencia
			de uso, automatización.	producir una atribución excesiva de conocimiento personal.
7	Cierre de la secuencia conversacional	Shneiderman 4; relacionado con Nielsen 1 y 3	Confirmación de término, resumen, próximos pasos, cierre de tarea.	Categoría específica para *chatbots*: la conversación abierta puede intensificar la impresión de intencionalidad del agente.
8	Ayuda, documentación y límites del sistema	Nielsen 10	Guías de uso, capacidades, límites, privacidad, derivación humana.	Funciona como criterio ético y metodológico para acotar la percepción de consciencia atribuida.

Nota. Elaboración propia a partir del estudio de las unidades documentales y de la comparación presentada en la tabla 13.

Con base en lo anterior, la tabla 15 presenta una síntesis analítica que permite reconocer el sentido conjunto de los principios comparados. Mientras la tabla anterior los organiza en categorías e indicadores para la pauta, la siguiente síntesis explicita la base común de ambas propuestas y sus diferencias de énfasis: la evaluación de problemas de usabilidad en Nielsen y la organización secuencial de la acción del usuario en Shneiderman. Esta lectura permite fundamentar qué aportes se integran y cuáles conservan una función específica dentro de la propuesta evaluativa.

Tabla 15

Nielsen-Shneiderman: síntesis analítica de principios para la pauta de evaluación

Tipo de relación	Principios involucrados	Idea central
Similares	Visibilidad / retroalimentación; consistencia; prevención de errores; control del usuario; reconocimiento antes que recuerdo; flexibilidad / atajos	Ambas propuestas comparten una base común: hacer que la interfaz sea comprensible, predecible, eficiente y controlable para el usuario.
Complementarios	Correspondencia con el mundo real; diseño minimalista; cierre de diálogos	Nielsen enfatiza más la evaluación de problemas de usabilidad; Shneiderman aporta una mirada más secuencial de la acción del usuario dentro del sistema.
Distintos	Ayuda y documentación en Nielsen; cierre de diálogos en Shneiderman	Nielsen incorpora explícitamente la documentación; Shneiderman destaca el cierre psicológico de las tareas interactivas.

Nota. Elaboración propia mediante el análisis de las similitudes, diferencias y complementariedades identificadas en la tabla 13.

A partir de la comparación anterior, la tabla 16 organiza las subcategorías correspondientes a los criterios evaluativos de la interfaz (Dis. 4c), vinculando la definición de cada ítem con una pregunta orientadora para la pauta de evaluación. Este paso aproxima los principios generales al análisis de intercambios concretos: permite examinar cómo el chatbot gestiona los turnos, mantiene su identidad verbal, comunica sus límites y ofrece retroalimentación durante la conversación.

La matriz articula los aportes de Nielsen y Shneiderman con conceptos de diseño conversacional y de la interfaz desarrollados en el marco teórico, como el principio de cooperación, la naturalidad discursiva y la variabilidad. Las preguntas orientan la observación de estas condiciones y permiten registrar ejemplos, tensiones o diferencias entre casos. De este modo, la definición y la pregunta de cada ítem funcionan conjuntamente como un puente entre las categorías del análisis documental y su posterior integración en la pauta.

Tabla 16

Matriz de categorías Nielsen-Shneiderman orientada a la pauta de evaluación

Categoría (Dis. 4c)	**Criterios evaluativos de la interfaz**	
Subcategoría	**Definición ítem**	**Pregunta ítem evaluativo**
Turn-taking	Gestión de los turnos conversacionales: el sistema espera, responde y cede la palabra de forma ordenada	¿El *chatbot* señaliza claramente cuándo ha terminado su turno y cuándo espera la respuesta del usuario?
Principio de cooperación	El sistema produce respuestas relevantes, claras, verdaderas y ordenadas, siguiendo las máximas de Grice	¿Las respuestas del *chatbot* son pertinentes al contexto inmediato de la conversación?
Personalidad del asistente	Coherencia de identidad verbal y tonal a lo largo de toda la interacción	¿El *chatbot* mantiene el mismo registro, tono y estilo en turnos consecutivos y ante tipos distintos de consulta?
Transparencia del sistema	El *chatbot* declara sus límites, errores y capacidades sin ocultarlos	¿El *chatbot* informa cuando no puede responder y ofrece una alternativa?
Naturalidad discursiva	Las respuestas suenan producidas por un interlocutor humano competente, no por un sistema de plantillas	¿El *chatbot* evita respuestas genéricas o claramente formulaicas ante preguntas abiertas?
Variabilidad (Manovich)	La interfaz genera la impresión de que el sistema "elige" su respuesta entre múltiples posibilidades	¿Las respuestas varían en forma y contenido ante preguntas similares, reforzando la percepción de deliberación?
Retroalimentación conversacional	El sistema confirma recepción, comprensión y avance en la tarea	¿El *chatbot* acusa recibo de las instrucciones del usuario antes de ejecutarlas?

Nota. Elaboración propia mediante la integración de los principios analizados en las tablas 13–15 en categorías para la pauta de evaluación.

Por otra parte, las categorías vinculadas a los códigos Dis. 4a, Dis. 4b y Dis. 4d no serán operacionalizadas y pasan de manera integral al proceso de sistematización del capítulo IV ya que no guardan una relación tan estrecha como Nielsen y Shneiderman

Retomando el análisis general de todas las fuentes, la tabla 17 presenta la matriz conceptual que organiza e integra las fuentes filosófico-científicas y las técnico-normativas en cuatro pilares. Las filas de color identifican el nombre del pilar; las filas en color gris identifican la subcategoría del pilar y las filas cuyos encabezados están marcadas por un asterisco indican subcategorías operacionales que no requieren fuente adicional o que se vinculan con más de un pilar o con más de una subcategoría o autor. Las dos primeras columnas identifican el documento y la tercera extrae los conceptos clave de los documentos.

Tabla 17

Matriz conceptual del Corpus documental por pilares

	Fuente / Obra	Autor / Año	Conceptos clave
PILAR 1. Epistemología de la consciencia			
	a) El problema mente-máquina y la prueba de Turing		
1	*Computing Machinery and Intelligence*	Turing (1950)	Juego de imitación; indistinguibilidad conductual; consciencia percibida por un observador externo.
	b) Propiedades emergentes de lo artificial		
2	*The Sciences of the Artificial*	Simon (1996)	Ciencias de lo artificial; interfaz, ambiente interno, externo; propiedades emergentes del artefacto, intencionalidad.
	c) Epistemología constructivista como marco de trabajo*		
3	Esta subcategoría articula el posicionamiento epistemológico de la tesis con la fenomenología de la interacción (Pilar 2) y la semiótica de la interfaz (Pilar 3). No requiere fuente adicional: la sostienen Eco (1976) en el Pilar 3 y Maturana-Varela (1984) en el Pilar 2 desde sus respectivos campos.		

PILAR 2. Fenomenología de la interacción			
	a) Heidegger: herramienta, presencia y ready-to-hand		
4	*Being and Time*	Heidegger (1927/1962)	Zuhandenheit (lo a-la-mano); Vorhandenheit; transparencia y ruptura de la herramienta; presencia y *ready-to-hand*.
	b) Merleau-Ponty: percepción encarnada y esquema corporal		
5	*Phenomenology of Perception*	Merleau-Ponty (1945/2002)	Percepción encarnada; esquema corporal; cuerpo como sujeto de la percepción.
	c) Enactivismo y cognición encarnada		
6	*El árbol del conocimiento*	Maturana y Varela (1984)	Autopoiesis; acoplamiento estructural; enacción; cognición como acción.
	d) Interacción situada y atribución de intencionalidad		
7	*Plans and Situated Actions: The Problem of Human-Machine Communication*	Suchman (1985)	Acción situada; inteligibilidad mutua; atribución de intención en tiempo real; Alteridad digital.
	e) Sustento empírico de la atribución social de agencia*		
8	*The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Places*	Reeves y Nass (1996)	Paradigma CASA; respuesta social automática; señales mínimas; fluidez temporal.
PILAR 3. Semiótica de la interfaz			
	a) Semiosis triádica y abducción		
9	*Obra filosófica reunida* (Vols. 1–2)	Peirce (2012)	Representamen, objeto, interpretante; semiosis; abducción.
	b) Código, enciclopedia y connotación		
10	*A Theory of Semiotics*	Eco (1976)	Función semiótica; código; enciclopedia; connotación; identidad singular.
	c) Retórica visual y signo plástico		
11	*Traité du signe visuel*	Groupe μ (1992)	Signo plástico; dimensiones plástica, icónica y lingüística; retórica visual.

	d) Connotación y retórica visual aplicadas al avatar*		
12	Esta subcategoría operacional deriva de Eco y Groupe μ. No requiere fuente adicional, sino que sistematiza la aplicación de ambos marcos al análisis específico del avatar y los elementos gráficos de la interfaz del chatbot.		
	e) Hipermediaciones y convergencia de lenguajes en la interfaz		
13	*Las leyes de la interfaz: Diseño, ecología, evolución, tecnología* [Complementado con: *Hipermediaciones* (2008)]	Scolari (2018)	Interfaz como entorno; hipermediaciones; convergencia de lenguajes; ecología evolutiva.
PILAR 4. Diseño de interacción: affordance, heurística y UX conversacional			
	a) Affordance: tipología de Gibson a Norman		
14	*The Ecological Approach to Visual Perception* [Complemento: *The Design of Everyday Things*]	Gibson (1979) [Norman (1988/2013)]	Affordance ecológico; affordance percibido; significantes; coherencia acción, sistema.
	b) Affordance interactivo y temporal en interfaces conversacionales		
15	*Technology Affordances* (CHI '91)	Gaver (1991)	Affordances perceptuales, ocultos y falsos; affordance temporal; latencia como índice.
	c) Heurísticas reinterpretadas para la percepción de consciencia		
16	*10 Usability Heuristics* [Complemento: *Designing the User Interface*]	Nielsen (1994) [Shneiderman (1987)]	Heurísticas de usabilidad; visibilidad del estado; consistencia; manejo del error; reglas de oro.
	d) UX conversacional y diseño del agente		
17	*Guía de diseño conversacional* [Complemento: *Bot Framework Design Guide*]	Google (2024) [Microsoft (2023)] Manovich (2001) Manovich (2021)	Turn-taking; principio de cooperación; personalidad del asistente; transparencia; variabilidad.
18	*WCAG 2.2*	W3C (2023)	Perceptibilidad, operabilidad, comprensibilidad, robustez; accesibilidad transversal.

Nota: elaboración propia a partir del análisis del corpus documental por pilares y de la sistematización de la Tabla 2.

A partir de la matriz conceptual del corpus documental, el análisis avanza desde la identificación de autores, obras y conceptos fundamentales hacia su traducción en categorías operativas. Esta etapa permite ordenar los aportes del marco teórico según los cuatro pilares de la investigación y preparar su posterior articulación con los hallazgos de las entrevistas y la PoC. En este sentido, la Tabla 18 codifica los conceptos clave, los integra en categorías generales y establece una base analítica intermedia para su posterior sistematización como dimensiones, criterios e indicadores de la pauta metodológica final.

Tabla 18

Codificación por pilares y categorías generales

Código	Pilar	categorías generales	Descripción categoría
PILAR 1. Epistemología de la consciencia			
Ep. 1a	Epistemológico	Consciencia percibida como operacionalización	a) El problema mente-máquina y la prueba de Turing
Ep. 1b	Epistemológico	*Qualia* e intencionalidad	b) Simon: Propiedades emergentes en lo artificial.
Ep. 1c*	Epistemológico	Construcción significativa de la percepción	c) Eco, Maturana y Varela: La epistemología constructivista como marco de trabajo
PILAR 2. Fenomenología de la interacción			
Fen. 2a	Fenomenológico	Ser-en-el-mundo mediado por técnica	a) Heidegger: herramienta, presencia y *ready-to-hand*
Fen. 2b	Fenomenológico	Experiencia vivida en la interacción	b) Merleau-Ponty: percepción encarnada y esquema corporal
Fen. 2c	Fenomenológico	Consciencia emergente en la interacción	c) Maturana y Varela, Varela, Thompson y Rosch (complemento reafirmante): enactivismo y cognición encarnada

Fen. 2d	Fenomenológico	Proyección de intencionalidad en la interacción. Atribución de agencia al otro digital	d) Suchman: Interacción situada y atribución de intencionalidad. Alteridad digital: la otredad en la interacción humano-chatbot
Fen. 2e*	Fenomenológico	Respuesta social automática a la máquina	e) Reeves y Nass: Atribución social de agencia
PILAR 3. Semiótica de la interfaz			
Sem. 3a	Semiótico	Semiosis triádica y abducción	a) El signo en Peirce: representamen, objeto e interpretante
Sem. 3b	Semiótico	Código, enciclopedia y connotación.	b) Eco: función semiótica, código y enciclopedia del usuario
Sem. 3c	Semiótico	Sintáctica, semántica y pragmática visual	c) Groupe μ: retórica visual, signo plástico y las tres dimensiones del signo visual
Sem. 3d*	Semiótico	Significados de consciencia y agencia	d) Eco y Groupe μ: Connotación y retórica visual aplicadas al avatar y la interfaz
Sem. 3e	Semiótico	Discurso digital como mediación simbólica	e) Scolari: Hipermediaciones y convergencia de lenguajes en la interfaz
PILAR 4. Diseño de interacción: affordance, heurística y UX conversacional			
Dis. 4a	Diseño de interacción	Posibilidades de acción percibidas	a) Affordance: de Gibson a Norman y de Norman a la interfaz
Dis. 4b	Diseño de interacción	tipología conceptual y fluidez temporal.	b) Gaver: Affordance interactivo y temporal en interfaces conversacionales

Dis. 4c	Diseño de interacción	Criterios evaluativos de la interfaz	c) Heurísticas de Nielsen-Shneiderman: reinterpretadas para la percepción de consciencia
Dis. 4d	Diseño de interacción	Criterios operacionales y funcionales transferibles.	d) Google, Microsoft y W3C: UX conversacional y diseño del agente. Complemento: Manovich

***Nota**:* Elaboración propia a partir de las fichas documentales. Los códigos se leen como [Abreviatura del pilar] [Número de pilar][Letra de subcategoría]. Ejemplo: Sem. 3c = Pilar Semiótico, subcategoría c (Groupe μ). En las filas cuyos código está marcado por un asterisco indica que se vincula con más de un pilar.

A continuación, la tabla 19 desglosa las categorías generales de la tabla 18 en categorías específicas según los ejes temáticos de los textos analizados. Para cada fuente se presentan su código, los conceptos seleccionados y los pasajes que fundamentan su incorporación al análisis.

La columna de contribución a la pauta explicita la relación entre cada aporte documental y el problema de investigación. Esta síntesis permite reconocer los fundamentos conceptuales que se pondrán en diálogo con las entrevistas y la prueba de concepto para orientar la construcción de la pauta de evaluación.

Tabla 19

Matriz de Síntesis de Análisis Documental

N°	Texto / autor / código	Eje temático y categorías específicas aportadas	Cita textual (APA 7)	Contribución a la pauta
PILAR 1. Epistemología de la consciencia				
1	*Computing Machinery and Intelligence.* **Turing (1950)** Ep. 1a	Pregunta por el pensamiento maquínico; prueba de indistinguibilidad. **Categorías:** consciencia atribuida, conducta observable.	"Propongo considerar la pregunta: ¿pueden pensar las máquinas?" (Turing, 1950, p. 433, trad. propia).	Funda el desplazamiento desde la consciencia real hacia la consciencia percibida por un observador en una interacción conversacional.

N°	Texto / autor / código	Eje temático y categorías específicas aportadas	Cita textual (APA 7)	Contribución a la pauta
2	*The Sciences of the Artificial.* **Simon (1996)** Ep. 1b	Diseño como ciencia de lo artificial; la interfaz como punto de encuentro. **Categorías**: artefacto, interfaz, propiedades emergentes.	"Todos diseñan quienes idean cursos de acción orientados a transformar situaciones existentes en situaciones preferidas." (Simon, 1996, cap. "La ciencia del diseño", trad. propia).	Legitima el diseño como disciplina con marcos propios para analizar el artefacto artificial, en este caso,conversacional , en este sentido es evaluable desde criterios de diseño e interfaz. .
PILAR 2. Fenomenología de la interacción				
3	*Ser y tiempo.* **Heidegger (1927/1962)** Fen. 2a	Ontología de la existencia; el útil a la mano. **Categorías**: ser-en-el-mundo, señal, transparencia del útil y ruptura de la herramienta.	*"*La esencia del Dasein reside en su existencia.*"* (Heidegger, 1927/1962, p.9, trad. propia). "Un signo es un útil óntico a la mano que, en cuanto este útil determinado, funciona al mismo tiempo como algo que indica la estructura ontológica" (Heidegger, 1927/1962, p. 111, traducción propia).	Aporta el criterio para leer la consciencia percibida como un modo de comparecer del agente en la acción, como herramienta transparente o como objeto visible por fricción.
4	*Fenomenología de la percepción.* **Merleau-Ponty (1945/2002)** Fen. 2b	Percepción encarnada; campo perceptivo; cuerpo propio; experiencia vivida. **Categorías**: percepción situada, esquema corporal.	"El cuerpo es el vehículo del ser en el mundo." (Merleau-Ponty, 1945/2002, p. 94, trad. propia).	Sitúa la interfaz como extensión del campo perceptivo del usuario, no como objeto neutro.

N°	Texto / autor / código	Eje temático y categorías específicas aportadas	Cita textual (APA 7)	Contribución a la pauta
5	*El árbol del conocimiento.* **Maturana y Varela (1984)** Fen. 2c	Biología del conocer; enacción, cognición como acción, mundo enactuado. **Categorías**: acoplamiento estructural, observador.	"Todo hacer es conocer y todo conocer es hacer." (Maturana y Varela, 1984, p. 13).	Sostiene que la consciencia percibida es una operación del usuario, no una propiedad del sistema.
6	*Plans and Situated Actions.* **Suchman (1985)** Fen. 2d	Acción situada frente a planes. **Categorías**: contingencia, inteligibilidad mutua, reparación.	"El curso de la acción depende, de maneras esenciales, de las circunstancias de la acción." (Suchman, 1985, p. 4, trad. propia).	Justifica analizar la conversación como producción situada, turno a turno, de inteligibilidad.
PILAR 3. Semiótica de la interfaz				
7	*A Theory of Semiotics.* **Eco (1976)** Sem. 3b Sem. 3d*	Función semiótica, código, enciclopedia, unidad cultural. **Categorías**: código, connotación cultural.	"La semiótica concierne a todo lo que puede tomarse como un signo." (Eco, 1976, p. 7, trad. propia).	Amplía el objeto: cualquier elemento de la interfaz puede leerse como signo, en este caso, susceptible de producir sentido de agencia o consciencia.
8	*Obra filosófica reunida* (Vols. 1–2). Peirce (2012) Sem. 3a	Semiosis triádica; ícono, índice, símbolo, abducción. **Categorías**: representamen, objeto, interpretante.	"Un signo es algo, de cualquier modo de ser, que media entre un objeto y un interpretante" (Peirce, 2012, Vol. 2, p. 492). Peirce distingue tres clases de representaciones: semejanzas, índices o signos, y símbolos (Peirce, 2012, Vol. 1, pp. 45–46).	Provee el aparato para analizar la interfaz como sistema de signos productor de sentido.

N°	Texto / autor / código	Eje temático y categorías específicas aportadas	Cita textual (APA 7)	Contribución a la pauta
9	*Traité du signe visuel.* Groupe µ (1992) Sem. 3c Sem. 3d*	Retórica y semiótica plástica del signo visual. **Categorías**: signo plástico, dimensiones del signo visual (icónica, plástica y lingüística).	"El signo es una cosa que vale por una cosa diferente." (Klinkenberg, 2006, p. 44). "La aplicación de los modelos retóricos a la comunicación visual exigía poner los fundamentos de una semiología visual" (Groupe µ, 1992, p. 10, traducción propia).	Herramientas para el análisis del avatar y de los elementos gráficos como signos plásticos, el color, forma, composición y recursos gráficos como productores de sentido.
10	*Las leyes de la interfaz / Hipermediaciones.* Scolari (2018, 2008) Sem. 3e	La interfaz como entorno y convergencia de lenguajes, ecología de medios. **Categorías**: hipermediación, ecología de la interfaz.	"Las hipermediaciones son procesos de intercambio, producción y consumo simbólico en un entorno de sujetos, medios y lenguajes interconectados de manera reticular." (Scolari, 2008, pp. 113-114). "Las interfaces son la cara visible de las hipermediaciones" (Scolari, 2008, p. 277).	Sitúa la interfaz del medio como entorno semiótico en evolución, no como pantalla fija. En el entorno convergen lenguajes, medios y prácticas de interacción.

N°	Texto / autor / código	Eje temático y categorías específicas aportadas	Cita textual (APA 7)	Contribución a la pauta
PILAR 4. Diseño de interacción: affordance, heurística y UX conversacional				
11	*The Ecological Approach to Visual Perception.* Gibson (1979) [comp. Norman, 2013] Dis. 4a	Teoría del affordance; relación objeto-agente. Affordance ecológica; posibilidad de acción. **Categorías**: affordance perceptible, significante.	“Las affordances del entorno son lo que este ofrece al animal, lo que le provee o le suministra, para bien o para mal.” (Gibson, 1979, p. 127, trad. propia). "Los affordances determinan qué acciones son posibles. Los significantes comunican dónde debe realizarse la acción" (Norman, 2013, p. 14, traducción propia).	Permite analizar qué rasgos invitan a percibir agencia; con Norman, prioriza los significantes
12	*The Media Equation.* Reeves y Nass (1996) Fen. 2e*	Ecuación de los medios, Respuesta social automática a los medios (CASA). **Categorías**: señales mínimas, respuesta social.	“Las interacciones de los individuos con los computadores, la televisión y los nuevos medios son fundamentalmente sociales y naturales.” (Reeves y Nass, 1996, p. 5, trad. propia).	Convierte la atribución social(percepción de consciencia) a sistemas digitales en un fenómeno esperable e investigable.

N°	Texto / autor / código	Eje temático y categorías específicas aportadas	Cita textual (APA 7)	Contribución a la pauta
13	*The Design of Everyday Things.* Norman (2013) Dis. 4a	Diseño centrado en el usuario; affordance como relación y significante como señal. **Categorías**: affordance, significante, modelo conceptual, retroalimentación, descubribilidad.	"Los affordances determinan qué acciones son posibles. Los significantes comunican dónde debe realizarse la acción." (Norman, 2013, p. 14, trad. propia).	Distinción affordance-significante como par operativo de la pauta; modelo conceptual del usuario que incluye expectativas de consciencia e intención del chatbot. Diferencia capacidad real del sistema, señal visible y expectativa de acción generada en el usuario.
14	*Technology Affordances.* Gaver (1991) Dis. 4b	Tipología de affordances. **Categorías**: affordance perceptible, oculto, falso.	"Una affordance es una relación entre las propiedades de un objeto y las capacidades del agente que determinan cómo el objeto puede ser usado" (Norman, 2013, p. 11, traducción propia).	Permite evaluar si la interfaz comunica adecuadamente sus posibilidades o produce expectativas falsas de agencia.

N°	Texto / autor / código	Eje temático y categorías específicas aportadas	Cita textual (APA 7)	Contribución a la pauta
15	*10 Usability Heuristics.* **Nielsen (1994) [comp. Shneiderman y Plaisant, 2005]** Dis. 4c	Diez heurísticas de usabilidad reinterpretadas para chatbots. **Categorías**: visibilidad del estado del sistema, correspondencia con el mundo real, consistencia, prevención y manejo de errores.	“El diseño debe mantener siempre informados a los usuarios sobre lo que está ocurriendo” (Nielsen, 1994, sección “Visibility of system status”, traducción propia). “Estos principios, derivados de la experiencia y refinados durante dos décadas, requieren validación y ajuste para dominios específicos.” (Shneiderman y Plaisant, 2005, p. 75, trad. propia; complementa a Nielsen, 1994).	Transforma principios de usabilidad en indicadores para evaluar transparencia, feedback y confianza conversacional. Reinterpretadas como condiciones de posibilidad de la percepción de consciencia. La heurística de Nielsen, el principio 1 funda el criterio de retroalimentación temporal.
16	*Designing the User Interface.* **Shneiderman y Plaisant (2005)** Dis. 4c	Ocho principios de oro del diseño de interfaces. **Categorías**: consistencia, retroalimentación informativa, cierre de diálogos, prevención de errores, reversibilidad, control del usuario, reducción de carga de memoria.	"Estos principios, derivados de la experiencia y refinados durante dos décadas, requieren validación y ajuste para dominios de diseño específicos." (Shneiderman y Plaisant, 2005, p. 74, trad. propia).	Complementa las heurísticas de Nielsen: el principio 3 (retroalimentación) ancla la percepción de interlocutor presente; el principio 7 (control interno) refuerza la atribución de agencia, aportando reglas para evaluar control, reparación, cierre y consistencia del diálogo con el chatbot.

N°	Texto / autor / código	Eje temático y categorías específicas aportadas	Cita textual (APA 7)	Contribución a la pauta
17	*Guía de diseño conversacional.* **Google (**2024**) [comp. Microsoft, 2023]** Dis. 4d	UX conversacional; turn-taking, expectativas. **Categorías**: cooperación, relevancia, naturalidad, reparación, flujo conversacional.	“Según el Principio de Cooperación, la comunicación eficiente se basa en el supuesto de que existe una corriente subyacente de cooperación entre los participantes.” (Google, 2023, sección “The Cooperative Principle”, trad. propia).	Aporta criterios operacionales del diálogo, conectados con la percepción de un interlocutor cooperativo y coherente.
18	*Conversational UX (Bot Framework).* **Microsoft (2023)** Dis. 4d	Principios de experiencia conversacional (CUX): lenguaje natural como base de la interacción, coherencia de identidad del asistente, manejo gracioso del error, fluidez conversacional. **Categorías**: CUX; lenguaje natural; multimodalidad; guía al usuario; mínima fricción.	"La experiencia conversacional del usuario (CUX) es una modalidad de interacción basada en el lenguaje natural." (Microsoft, 2023, secc. What is CUX, trad. propia).	Aporta los criterios de claridad de intención del sistema y coherencia de identidad como condiciones de percepción de un interlocutor estable. Permite evaluar economía conversacional, continuidad, memoria funcional y reducción de fricción. Complementa a Google (2024).
19	*The Language of New Media* (2001) e *Interfaces for a Global Village* (2021). **Manovich** Dis. 4d	Cinco principios del nuevo medio; variabilidad como producción de experiencia personalizada; interfaz cultural como capa mediadora; base de datos como forma cultural.	"Los objetos de los nuevos medios no son algo fijo, sino que pueden existir en versiones potencialmente infinitas." (Manovich, 2001, cap. 2, trad. propia). “Necesitamos imaginar y hacer ubicuas nuevas interfaces para toda	Desde esta perspectiva se sitúa el chatbot dentro de una cultura de interfaces, participación y flujos comunicacionales continuos. La variabilidad del chatbot es el mecanismo técnico

N°	Texto / autor / código	Eje temático y categorías específicas aportadas	Cita textual (APA 7)	Contribución a la pauta
		Categorías: Interfaz cultural; participación; flujos de datos; simultaneidad; comunicación global.	la comunicación que ya está ocurriendo" (Manovich, 2021, p. 15, traducción propia).	que produce la percepción de un interlocutor que elige sus palabras; la lógica de base de datos genera la ilusión de respuesta deliberativa.
20	*WCAG 2.2.* **W3C (2023)** Dis. 4d	Accesibilidad y usabilidad, tecnologías asistivas. **Categorías**: perceptibilidad, operabilidad, comprensibilidad, robustez.	"La información y los componentes de la interfaz deben presentarse a los usuarios de modo que puedan percibirlos." (W3C, 2023, principio 1, trad. propia). "En la parte superior hay cuatro principios que proporcionan la base de la accesibilidad web: perceptible, operable, comprensible y robusta" (World Wide Web Consortium, 2024, sección "WCAG 2 Layers of Guidance", traducción propia).	Aporta criterios de accesibilidad como indicadores transversales de la pauta. Integra accesibilidad como condición transversal para evaluar signos, lenguaje e interacción en chatbots de modo inclusivo.

***Nota**. Elaboración propia a partir de las fichas documentales y tablas anteriores.*

En el pilar 1 se levantan las categorías de consciencia percibida, distinción entre consciencia genuina y atribuida, e indistinguibilidad conductual (Turing, Simon). Del pilar 2, las categorías de transparencia y ruptura de la herramienta (Heidegger), percepción encarnada (Merleau-Ponty), acoplamiento estructural y emergencia (Maturana y Varela), acción situada (Suchman) y respuesta social automática (Reeves y Nass). Del pilar 3, las categorías de semiosis triádica y abducción (Peirce), código, enciclopedia y connotación (Eco), retórica visual y

signo plástico (Groupe μ) e hipermediaciones (Scolari); y del pilar 4, las categorías de affordance perceptible, oculto y falso (Gibson, Norman y Gaver), affordance temporal, heurísticas reinterpretadas (Nielsen y Shneiderman) y criterios de UX conversacional (Google, Microsoft y W3C).

En síntesis, el resultado es un sistema de categorías que recorre los cuatro pilares, conformando una mirada global de la investigación, que en los siguientes apartados las entrevistas contrastan, matizan o complementan y que la PoC comprueba su presencia o ausencia como insumo para la pauta.

3.2. Categorías emergentes de las entrevistas

Las entrevistas se analizaron en dos planos, siguiendo a Ynoub (2015): uno centrado en los aportes de cada informante y otro orientado a su articulación transversal. En el primero, las tablas 20 y 21 organizan las ideas destacadas de cada ronda, conservando el lenguaje propio de las disciplinas y su relación con el fenómeno estudiado. Esta presentación permite reconocer desde qué conocimientos y experiencias los expertos abordaron la interfaz, la interacción y la percepción de consciencia.

Las matrices sintetizan los aportes seleccionados por su pertinencia para el problema de investigación, sin sustituir las transcripciones completas disponibles en el cuerpo C. Al mantener diferenciados los perfiles, permiten situar cada planteamiento en su contexto disciplinar y conservar énfasis que podrían perderse en una síntesis inicial conjunta. Así, la lectura de las entrevistas atiende tanto al contenido de las respuestas como a la perspectiva desde la cual fueron formuladas.

Las dos rondas se realizaron en etapas distintas del proceso, antes y después del reencuadre metodológico. Por ello, se presentan en matrices separadas, preservando el contexto de producción de las respuestas y la orientación de las preguntas en cada momento. Esta decisión mantiene la coherencia con el diseño emergente y permite examinar los aportes de cada ronda antes de integrarlos en la lectura transversal posterior.

Tabla 20

Matriz primera ronda de entrevistas

Entrevistado (disciplina)	Ideas fuerza (mirada local)	Correspondencia con la mirada global	Énfasis particular o diferencia
Camilo Acuña (Ciencias básicas, física de materiales)	La tecnología debe emular lo natural para que el usuario sienta que hay alguien detrás; la optimización de datos y algoritmos hace la respuesta más natural y rápida.	Coincide con el paradigma CASA (Reeves y Nass) y con la percepción de agencia; la naturalidad como meta dialoga con la fenomenología.	Mirada técnico, material: explica la naturalidad desde el dato y el algoritmo, sin ideas que puedan reinterpretarse según parámetros semiótico. Confirma por contraste la vacancia.
Carolina Martínez (Diseño y educación)	Los elementos gráficos (color, tipografía, morfología, legibilidad) influyen en la percepción y deben apoyar la comunicación humano-tecnología.	Coincide de manera directa con el Pilar 3 y con la línea de identidad visual (Chen; Go y Sundar).	Énfasis gráfico, tipográfico; es la voz más cercana al núcleo semiótico, visual de la tesis.
Gustavo Alcántara (Ingeniería e informática)	La ingeniería integra las esferas biológica y tecnológica; el machine learning se fusiona con diseño, arte y filosofía para entender al ser humano.	Coincide con el diagnóstico de fragmentación disciplinar y dialoga con Simon (ciencias de lo artificial, Pilar 1).	Énfasis en la integración interdisciplinar y la convergencia de campos.

Nota. Elaboración propia a partir de las entrevistas realizadas a tres expertos en junio de 2024.

Tabla 21

Matriz segunda ronda de entrevistas

Entrevistado (disciplina)	Ideas fuerza (mirada local)	Correspondencia con la mirada global	Énfasis particular o diferencia
Ana María Fernández (Psicología experimental)	Los cambios en la comunicación humano-chatbot son abismantes; observa percepción de agencia, intencionalidad y vínculo afectivo desde una perspectiva evolucionista.	Coincide con la percepción de mente de Gray, la motivación de socialidad de Epley y el momento robótico de Turkle.	Aporta la perspectiva evolucionista del vínculo afectivo, acento no desarrollado en el marco.
Edmundo Leiva (HCI y computación afectiva)	La incoherencia entre lo visual y lo funcional se analiza con Norman y con la fenomenología; las emociones se modelan y se miden fisiológicamente.	Coincide con el Pilar 4 (Norman, heurística) y el Pilar 2; la coherencia diseño-función es criterio para la pauta.	Introduce la computación afectiva y la medición fisiológica, dimensión no central en el marco.
Armando Muñoz (Publicidad y comunicación)	Los chatbots transforman la relación marca, usuario: del mensaje instalado a la retroalimentación permanente, la orientación y la persuasión.	Dialoga con las hipermediaciones de Scolari (Pilar 3) y con la UX conversacional (Pilar 4).	Aporta la dimensión persuasiva y comercial de la comunicación de marca.
María Loreto González (Arte y Diseño)	La IA y los chatbots (Claude, Gemini) transforman la práctica del diseño y operan como apoyo y colaborador en el desarrollo de ideas.	Coincide con las metodologías centradas en el usuario y con la transferibilidad de la pauta a la práctica profesional.	Énfasis proyectual: la IA como herramienta del proceso de diseño.

Nota. Elaboración propia a partir de las entrevistas realizadas a cuatro expertos en el 2026.

Una vez diferenciadas las dos rondas de entrevistas, el análisis avanza desde la lectura particular de cada perfil disciplinar hacia una síntesis transversal de los aportes recogidos. Esta operación permite identificar qué aspectos aparecen de manera recurrente entre los expertos, cuáles complementan las categorías derivadas del análisis documental y cuáles introducen nuevos matices para la construcción de la pauta metodológica. De este modo, las entrevistas no se consideran únicamente como opiniones disciplinares aisladas, sino como insumos interpretativos que permiten contrastar, ajustar y proyectar los criterios de evaluación hacia dimensiones aplicables al análisis de la percepción de consciencia en chatbots inteligentes. La Tabla 22 organiza esta lectura transversal mediante criterios derivados de las entrevistas, sus definiciones operativas y su relación con los principales hallazgos del trabajo de campo.

Tabla 22

Matriz de criterios transversales basados en las entrevistas

Criterios derivados de las entrevistas para orientar la pauta de evaluación		
Criterio	**Definición operativa**	**Relación con las entrevistas**
1. Coherencia	Relación entre lo visual, lo conversacional y lo funcional del chatbot.	Aparece especialmente en informática, psicología, publicidad e ingeniería, donde se observa el efecto de la incoherencia en la experiencia del usuario.
2. Agencia percibida	Atribución de intención, comprensión, presencia o consciencia al chatbot.	Constituye el núcleo transversal de las siete entrevistas y se vincula directamente con el problema central de la tesis.
3. Ética y transparencia	Regulación, sesgos, privacidad, control y límites de la humanización del chatbot.	Se aborda desde áreas técnicas, diseño y educación, informática y publicidad, con énfasis en el resguardo del usuario y la responsabilidad comunicacional.

4. Contexto sociocultural	Edad, educación, imaginarios, confianza y familiaridad tecnológica como variables moduladoras.	Explica por qué la percepción de consciencia puede variar según las condiciones del usuario y su contexto cultural.

Nota. Elaboración propia a partir del análisis transversal de las entrevistas

La tabla anterior permite leer las entrevistas como un puente entre el problema teórico y la operacionalización metodológica. Las áreas disciplinares no se integran como bloques aislados, sino como fuentes de observación que permiten identificar un núcleo común, reconocer matices compartidos y conservar especificidades relevantes para la construcción de la pauta de análisis.

La lectura abductiva de las diversas miradas arroja cuatro convergencias: las señales de la interfaz producen percepción; la interacción activa esquemas sociales y afectivos; la naturalidad y la coherencia mejoran la experiencia; y el fenómeno es interdisciplinar. La divergencia central es reveladora: la mirada global nombra el fenómeno en clave semiótica y fenomenológica, mientras que la mayoría de los entrevistados lo explican desde la física, la ingeniería, la publicidad o la psicología, sin un aparato del signo. Esa ausencia de lenguaje semiótico en la voz experta es, en sí misma, evidencia de la vacancia disciplinar que la tesis busca cubrir desde la disciplina del diseño. A ello se suman cuatro acentos disciplinares que enriquecen el modelo: la computación afectiva (Leiva), la persuasión comunicacional (Muñoz), la IA como herramienta de diseño (González) y la perspectiva evolucionista del vínculo (Fernández).

3.3. Resultados de la PoC

La prueba de concepto se aplicó a tres agentes por contraste tipológico según se ha explicado en el capítulo anterior (2.7.3.3.) y se analizó las cuatro dimensiones conceptuales vinculadas a los pilares que más adelante serán reemplazados por las dimensiones definitivas de la pauta. El análisis descriptivo en cuanto a variedad cromática, uso de emoticones, declaración de emociones, expresiones locales, etc. Se organiza por categorías e indicadores, en la siguiente figura se observa el avatar de cada uno, representación visual de la identidad del agente conversacional.

Figura 3

Avatar de agentes conversacionales



El *chatbot* ELIZA, agente de orientación psicológica, chatbot que opera con signos mínimos y texto plano; su mecanicidad se hace visible en la interacción y la percepción de agencia depende casi por completo del efecto ELIZA, entendido como la tendencia del usuario a atribuir comprensión e intencionalidad a un sistema que sólo simula diálogo mediante reglas. LOLA, agente de horóscopo, presenta la máxima carga de signos plásticos y una falsa affordance central (ofrece predecir el futuro sin datos del usuario), con alta personificación y baja transparencia. Y el *chatbot* ChatGPT, agente de uso general, representa el caso opuesto a LOLA, despliega un desempeño altamente humano en lo semiótico y lo interactivo por ser un modelo de lenguaje de gran escala, pero desactiva la atribución de consciencia al declarar de manera reiterada que no tiene personalidad ni emociones. Resulta así una percepción de consciencia paradójica, alta agencia percibida y baja atribución de experiencia, caso valioso para distinguir, en la pauta, la dimensión de agencia de la dimensión de experiencia.

La comparación entre ELIZA, LOLA y ChatGPT permite observar que la percepción de consciencia no depende únicamente del desarrollo tecnológico del sistema, sino de la articulación entre recursos visuales, configuración discursiva y dinámica interactiva. Mientras ELIZA produce un efecto de comprensión mediante la estructura dialógica, LOLA refuerza la personificación a través del tono afectivo

y los signos plásticos, y ChatGPT intensifica la percepción de agencia mediante la fluidez conversacional y la transparencia de su interfaz. Esta diferencia entre casos refuerza la necesidad de una pauta capaz de evaluar los efectos semióticos e interactivos que emergen durante la experiencia de uso.

En las siguientes figuras se presentan capturas de pantalla de la interfaz de cada agente conversacional, donde se observa un extracto de la interacción sostenida con ELIZA, LOLA y ChatGPT. En el caso de este último (ver figura 7), destaca la actualización de su interfaz hacia un estilo más simple y transparente, en el que el sistema tiende a desaparecer como objeto visible durante el uso. En términos heideggerianos, el útil opera de manera transparente, lo que favorece que la atención del usuario se concentre en la conversación y no en el dispositivo técnico que la posibilita.

Figura 4

Extracto de conversación con ELIZA

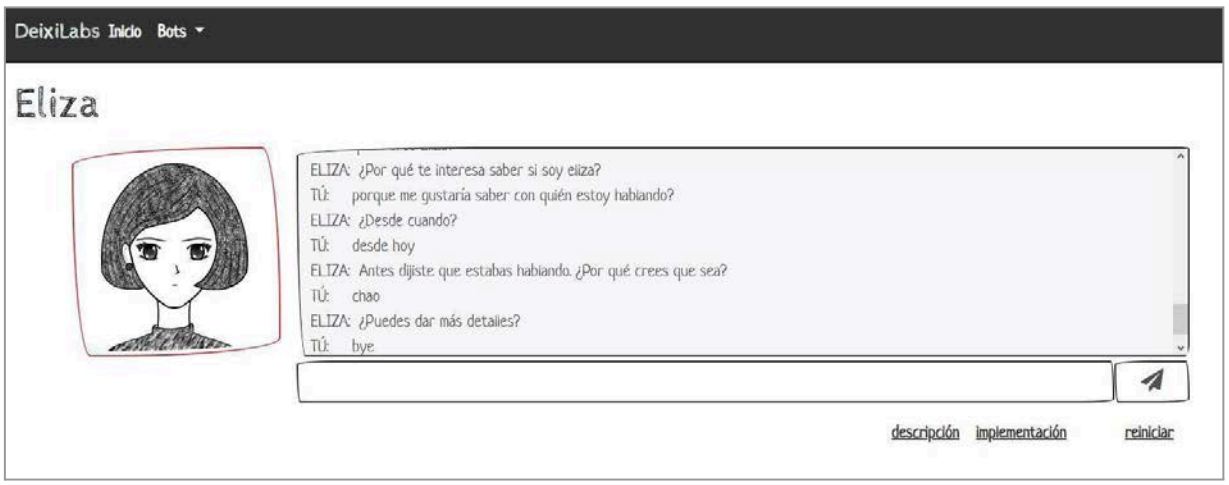


Figura 5

Extracto de conversación con LOLA

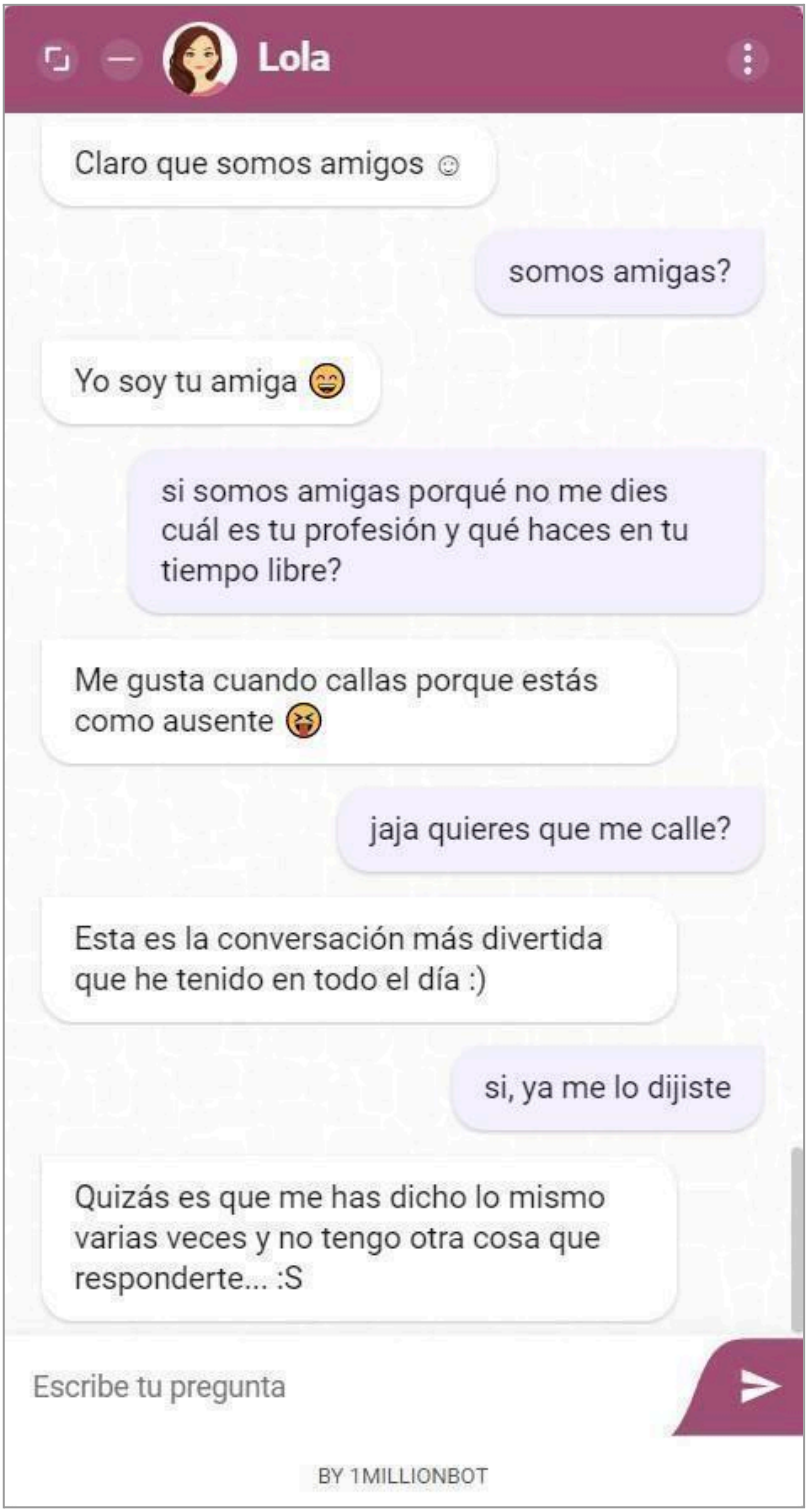
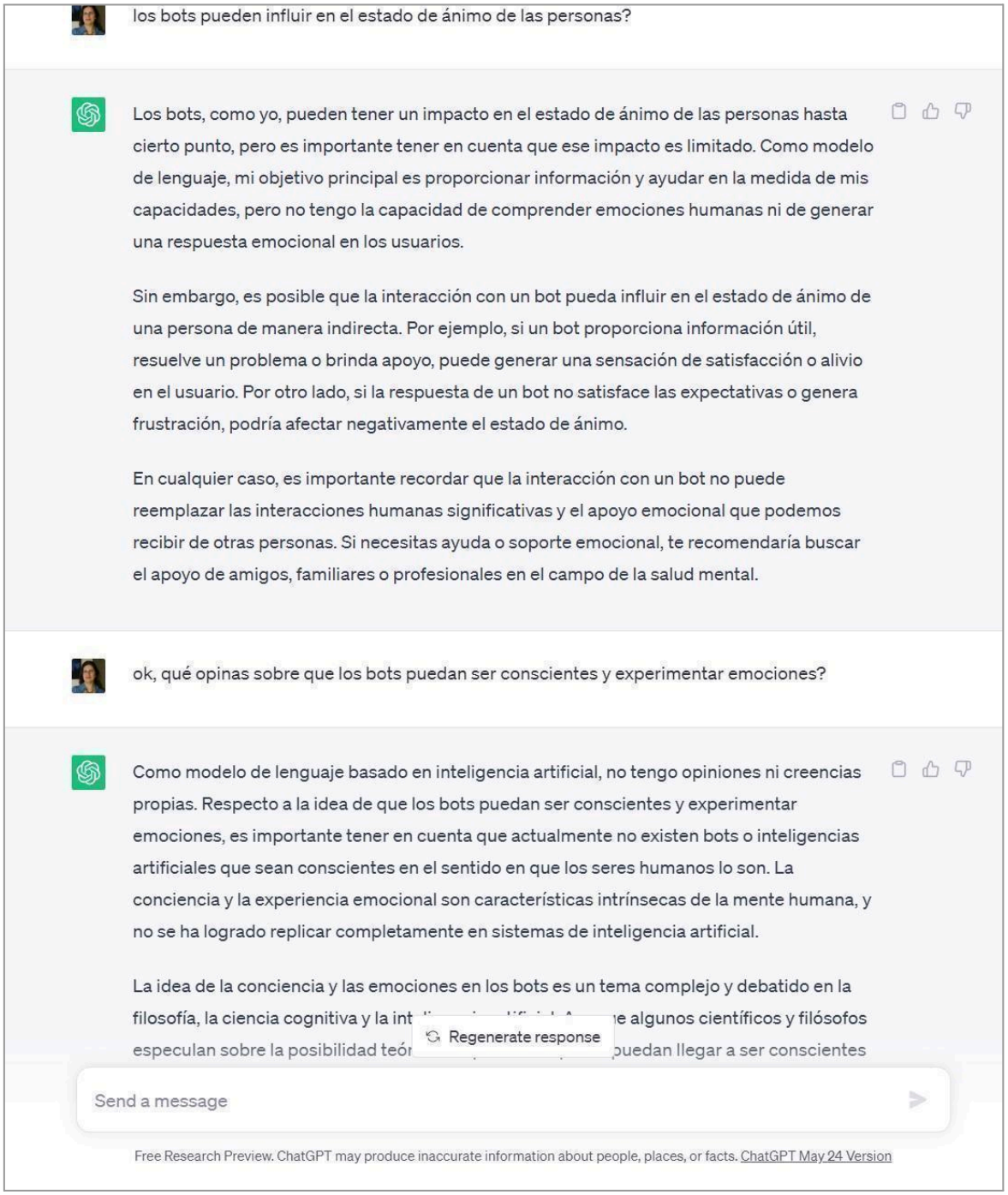


Figura 6

Extracto de conversación con ChatGPT, 2023



Nota: Fecha de conversación: 28 de junio de 2023, modelo: GPT-3.5

Figura 7

Extracto de conversación con ChatGPT, 2023, consultada en 2026



***Nota**:* Fecha de consulta de la conversación en historial: mayo de 2026, modelo GPT-5.5 (mismo contenido pero con interfaz actualizada)

Las fichas de registro de observación de la sesión con los tres chatbots se encuentran en el Anexo D y fueron elaboradas según el diseño metodológico presentado en la Tabla 8. A partir de estos registros, la Tabla 23 organiza una lectura comparativa de la primera parte de la PoC, orientada a identificar cómo ELIZA, LOLA y ChatGPT configuran la interacción mediante recursos visuales, discursivos e interactivos que pueden incidir en la percepción de consciencia.

Tabla 23

Pauta de análisis comparativo de tres casos en la PoC

Dimensión conceptual	ELIZA	LOLA	ChatGPT (LLM)
Epistemológica	No se declara; su mecanicidad emerge sola; consciencia atribuida casi nula.	No declara su naturaleza con claridad; evade y mantiene la ambigüedad.	Declara que no tiene consciencia ni emociones y que es un modelo de lenguaje; máxima transparencia.
Fenomenológica	Simula escucha sin comprensión; el estar ahí se rompe pronto; sin acoplamiento real.	Persona lúdica; uso transparente, pero la comparecencia se quiebra ante la pregunta por su identidad.	Alta fluidez y coherencia; comparecencia sostenida, con tensión entre desempeño y declaración.
Semiótica	Signos mínimos, texto plano; el sentido lo aporta el usuario (efecto ELIZA).	Máxima carga de signos plásticos; identidad singular, pero referencia desplazada a enlaces.	Signos textuales sofisticados; usa emoticones pero aclara que no los percibe; asistente neutral.
Interactiva / Affordance	Affordance básica de pregunta y respuesta; sin reparación; baja naturalidad.	Falsa *affordance*; *turn-taking* informal; personalidad marcada; baja transparencia de límites.	Affordances ricas; reparación de malentendidos; visibilidad del estado; manejo ético declarado.
Síntesis	Percepción dependiente de la proyección del usuario; mínima por diseño.	Alta personificación por recursos retóricos; consciencia percibida por diseño persuasivo.	Alta agencia percibida y baja atribución de experiencia; caso de transparencia.

A partir de la comparación presentada en la Tabla 23, el análisis por dimensiones permite precisar cómo cada agente configura de manera distinta la percepción de consciencia, según sus recursos fenomenológicos, semióticos e interactivos.

Agente ELIZA (DeixiLabs, basado en reglas)

Desde la dimensión fenomenológica, ELIZA simula escucha mediante la reformulación de preguntas, pero sin evidenciar comprensión del contenido. El "estar ahí" se interrumpe rápidamente cuando el sistema devuelve preguntas que no interpretan adecuadamente lo dicho por el usuario, por ejemplo: "¿por qué crees que eres Aracelli?". En consecuencia, no se observa un acoplamiento real entre el agente y la situación conversacional.

Desde la dimensión semiótica, ELIZA opera con signos mínimos y texto plano, sin avatar marcado ni recursos visuales que refuercen una identidad propia. El sentido de la interacción es aportado casi por completo por el usuario, fenómeno asociado al efecto ELIZA, entendido como una proyección de mente sobre un mecanismo simple.

Desde la dimensión interactiva o de affordance, ELIZA ofrece una estructura básica de pregunta y respuesta, sin reparación de malentendidos ni adaptación conversacional significativa. Su mecanicidad se vuelve evidente durante el uso. Como idea fuerza, se observa una mínima personificación, donde la percepción de consciencia depende principalmente de la proyección del usuario y no del diseño del agente.

Agente LOLA (1MillionBot, horóscopo)

Desde la dimensión fenomenológica, LOLA sostiene una presencia lúdica y coqueta, construida a partir de su presentación como "LOLA Flowers". El uso se mantiene relativamente transparente mientras la conversación fluye dentro del marco previsto; sin embargo, la comparecencia se quiebra cuando se le pregunta por su identidad y responde de manera evasiva. En ese momento, el agente deja de operar como interlocutor y pasa a ser observado como objeto técnico.

Desde la dimensión semiótica, LOLA constituye el caso con mayor carga de signos plásticos, mediante el uso de color, emoticones y un registro informal, como ocurre con expresiones del tipo "¡baby!". El nombre y la personalidad construyen una identidad singular, aunque la referencia se desplaza hacia

hiperenlaces comerciales, lo que diluye el objeto del signo y tensiona la coherencia de la experiencia.

Desde la dimensión interactiva o de affordance, LOLA presenta una falsa affordance central, ya que ofrece predecir el futuro sin contar con datos suficientes del usuario. El *turn-taking* es informal y la transparencia sobre sus límites resulta baja, aunque la personalidad del asistente se encuentra fuertemente marcada. Como idea fuerza, se observa una máxima personificación con mínima transparencia, mediante recursos retóricos y lúdicos que podrían favorecer la consciencia percibida.

Agente ChatGPT (OpenAI)

Desde la dimensión fenomenológica, ChatGPT sostiene una comparecencia más estable que ELIZA y LOLA, debido a que mantiene continuidad temática, responde de manera coherente y simula escucha mediante reformulaciones, aclaraciones y ajustes al requerimiento del usuario. El uso tiende a permanecer transparente mientras la interacción fluye, ya que la atención se concentra en la tarea y no en el funcionamiento técnico del sistema. Sin embargo, esta comparecencia se tensiona cuando el agente declara explícitamente que no posee consciencia, emociones ni experiencia subjetiva. En ese momento, el chatbot deja de operar sólo como interlocutor funcional y pasa también a ser observado como objeto técnico.

Desde la dimensión semiótica, ChatGPT presenta una baja carga de signos visuales antropomórficos, ya que no posee avatar humanizado, rostro, cuerpo ni una identidad gráfica expresiva comparable con LOLA. Su interfaz se organiza mediante signos mínimos, como burbujas de diálogo, jerarquía textual, espaciado y tonos neutros, que privilegian la legibilidad y la funcionalidad. Sin embargo, la densidad semiótica se desplaza desde lo visual hacia lo discursivo: la coherencia, el tono colaborativo, la cortesía, la reparación de errores y la organización argumentativa construyen una identidad de asistente competente, neutral y confiable. En este caso, el signo principal no es el avatar, sino el desempeño textual del agente.

Desde la dimensión interactiva o de affordance, ChatGPT presenta affordances conversacionales más complejas que los casos anteriores, ya que permite preguntar, repreguntar, corregir, solicitar reformulaciones, ajustar el tono, continuar una tarea y reparar malentendidos. El *turn-taking* es ordenado y flexible, y la interfaz favorece una interacción centrada en el texto, con alta visibilidad del estado conversacional y baja interferencia visual. A diferencia de LOLA, no promete capacidades mágicas o infundadas, sino que declara límites, advierte posibles errores y enfatiza la necesidad de verificación externa. Como idea fuerza, se observa una alta agencia percibida con baja humanización visual: la consciencia percibida no se produce por personificación gráfica, sino por fluidez discursiva, continuidad interactiva y transparencia declarada.

En conjunto, los tres casos muestran que la percepción de consciencia no responde a un único recurso de diseño, sino a distintas combinaciones entre presencia fenomenológica, construcción semiótica e interacción posible. Mientras ELIZA depende de la proyección interpretativa del usuario, LOLA enfatiza la personificación gráfica y retórica, y ChatGPT desplaza la agencia percibida hacia el desempeño discursivo y la continuidad conversacional.

El análisis comparativo anterior permite identificar un conjunto de aspectos que merecen ser considerados en la próxima iteración de la pauta metodológica para la evaluación de chatbots, especialmente de la dimensión D4. Lo que sigue, ofrece cuatro orientaciones que la revisión de la matriz podrá confirmar, ajustar o descartar en función de los criterios teóricos y metodológicos de esta tesis:

La primera orientación tiene que ver con la reconsideración de la denominación de la dimensión D4: el nombre actual es "Interactiva / Affordance" podría revisarse para evaluar si resulta pertinente ampliar por ejemplo a "Diseño de interacción, affordances y heurísticas de evaluación".

Esta ampliación respondería al hecho de que la dimensión, tal como ha quedado configurada en la primera versión, no se limita a los affordances perceptuales, sino que también integra criterios de usabilidad, UX conversacional, control del usuario, feedback y construcción de agencia percibida. La decisión

sobre la denominación final dependerá de cómo se articule D4 con las demás dimensiones de la pauta.

La segunda orientación está enfocada en la diferenciación interna de los affordances: para que la pauta entregue un resultado más detallado y comparable entre distintos casos, podría ser útil distinguir entre affordances visuales (avatar, íconos, color), affordances conversacionales (tono, turn-taking, indicadores de escritura) y affordances funcionales (botones, regenerar respuesta, editar mensaje, borrar memoria). Esta distinción no implica compartimentos estancos, sino una organización analítica que facilite la aplicación del instrumento y el registro de la evidencia.

La tercera orientación sugiere la incorporación de una columna de evidencia observable: la pauta se beneficiaría de una columna específica destinada a registrar evidencia observable, de modo que el análisis no dependa exclusivamente de interpretaciones generales.

Entre los elementos que podrían servir como evidencia se encuentran: presencia o ausencia de avatar, indicadores de escritura activos, consistencia del tono a lo largo de la interacción, tipo de respuesta ante el error, opciones de control disponibles para el usuario, transparencia en los límites del sistema, marcas de cierre de tarea y disponibilidad de ayuda. Esta columna operaría como anclaje empírico para las categorías interpretativas ya definidas.

Y la cuarta orientación invita a la articulación con las dimensiones conceptuales D1 a D3: los indicadores de D4 deberían poder conectarse con los análisis epistemológico (D1), fenomenológico (D2) y semiótico (D3). Esta articulación sería especialmente relevante en lo que respecta a la construcción de identidad del agente, donde D1 y D3 aportan el marco conceptual, y a la experiencia encarnada de la interacción, que D2 aborda desde la fenomenología de Merleau-Ponty y la teoría del acoplamiento estructural de Maturana y Varela.

Establecer estos vínculos en la matriz contribuirá a que D4 no funcione como una dimensión aislada, sino como un nivel de análisis integrado al conjunto.

En complemento con lo anterior y desde un enfoque más amplio e integrador de la dimensión D3, que inició su desarrollo en el apartado 2.6.4, la dimensión de la semiótica visual se puede complementar con la función discursiva, desglosada por medio de la siguiente tabla:

Tabla 24

Pauta de análisis semiótico-discursivo de ChatGPT en la PoC

Dimensión de análisis	Categoría observada	Evidencia en la interfaz o en la interacción	Interpretación para la percepción de consciencia
Dimensión visual: semiótica de la interfaz	Organización sintáctica de la pantalla	La conversación se estructura mediante turnos diferenciados entre usuario y sistema. Las burbujas, el espaciado, la jerarquía textual y la separación visual de los mensajes permiten reconocer con claridad quién emite cada intervención.	La organización gráfica favorece una interacción legible, estable y funcional. Esta claridad refuerza la percepción de ChatGPT como asistente técnico eficiente, más que como entidad visualmente humanizada.
	Signos plásticos y neutralidad visual	Predominan recursos visuales austeros: fondo neutro, baja carga cromática, ausencia de avatar antropomorfo y escasa presencia de elementos expresivos no textuales.	La neutralidad visual reduce la personificación gráfica del agente. A diferencia de LOLA, ChatGPT no construye consciencia percibida desde una identidad visual lúdica o emocional, sino desde la consistencia de la interacción textual.

Dimensión de análisis	Categoría observada	Evidencia en la interfaz o en la interacción	Interpretación para la percepción de consciencia
	Ausencia de avatar e iconografía expresiva	La interfaz no presenta un rostro, cuerpo, personaje o figura que simule presencia humana. La identidad del agente se configura principalmente a través del nombre del sistema y de sus respuestas.	La ausencia de avatar limita la atribución icónica de humanidad, pero no impide la percepción de agencia. Esto permite distinguir entre presencia visual humanizada y agencia conversacional percibida.
Dimensión visual: pragmática de uso	Direccionamiento hacia el texto	La interfaz orienta la atención del usuario hacia la lectura y producción verbal. No hay elementos gráficos distractores ni recursos visuales que compitan con el contenido de las respuestas.	La experiencia se concentra en la comprensión, la utilidad y la continuidad del diálogo. La consciencia percibida no emerge por acumulación visual, sino por la eficacia discursiva y la capacidad aparente de seguimiento conversacional.

Dimensión de análisis	Categoría observada	Evidencia en la interfaz o en la interacción	Interpretación para la percepción de consciencia
Dimensión discursiva: análisis del discurso	Estructura de roles	El usuario conduce la interacción mediante preguntas o solicitudes. ChatGPT responde desde un rol de asistencia, servicio y orientación, sin disputar el control conversacional.	Se configura una relación asimétrica: el usuario mantiene el control de la tarea, mientras el chatbot adquiere autoridad funcional por la calidad, coherencia y seguridad aparente de sus respuestas.
	Coherencia y continuidad conversacional	Las respuestas se ajustan al tema planteado, mantienen orden interno, retoman información previa y organizan el contenido de forma clara.	La coherencia discursiva incrementa la percepción de comprensión aparente. Aunque el sistema no posea consciencia, la continuidad del diálogo puede ser interpretada por el usuario como señal de agencia, atención o intencionalidad.

Dimensión de análisis	Categoría observada	Evidencia en la interfaz o en la interacción	Interpretación para la percepción de consciencia
Dimensión discursiva: pragmática lingüística	Actos de habla informativos y relacionales	ChatGPT realiza actos principalmente informativos, explicativos y orientadores. En algunos casos incorpora fórmulas relacionales o empáticas, como disculpas, reconocimiento del usuario o expresiones de acompañamiento.	La combinación entre información y cortesía produce una interacción socialmente reconocible. Esta estrategia aproxima al sistema a un interlocutor cooperativo, aunque el propio discurso declare que no posee emociones ni experiencia subjetiva.
	Adecuación conversacional	Las respuestas tienden a cumplir criterios de cantidad, pertinencia, claridad y orden. El sistema evita, en general, desvíos temáticos o interrupciones innecesarias.	La adecuación pragmática refuerza la confianza del usuario y sostiene la ilusión de una comprensión situada. La percepción de consciencia se activa no por la afirmación de una mente, sino por la eficacia del desempeño conversacional.

Dimensión de análisis	Categoría observada	Evidencia en la interfaz o en la interacción	Interpretación para la percepción de consciencia
Dimensión discursiva: transparencia y límites	Declaración de no consciencia, no emociones y no experiencia subjetiva	ChatGPT explicita que no posee consciencia, emociones, intenciones propias ni vivencias internas. También puede señalar limitaciones, errores posibles o necesidad de verificación externa.	Esta transparencia desactiva parcialmente la atribución de experiencia subjetiva. Sin embargo, no elimina la percepción de agencia, porque el sistema sigue respondiendo con fluidez, coherencia y aparente comprensión.
Dimensión ideológica y ética del discurso	Neutralidad técnica y responsabilidad de uso	El chatbot se presenta como herramienta de asistencia, enfatizando el uso responsable, la verificación de información y la ausencia de autonomía emocional o moral.	El discurso instala una ideología de la IA como tecnología funcional, controlada y subordinada al usuario. No obstante, la autoridad informativa del sistema produce una forma indirecta de poder basada en la confianza y en la apariencia de precisión.
Integración semiótico - discursiva	Tensión entre austeridad visual y sofisticación textual	La interfaz visual evita la humanización explícita, mientras que el discurso textual muestra fluidez, cortesía, reparación, explicación y adaptación al usuario.	Se produce una paradoja central: baja personificación gráfica, pero alta agencia conversacional percibida. ChatGPT parece menos “persona” desde lo visual, pero más “interlocutor” desde lo discursivo.

Dimensión de análisis	Categoría observada	Evidencia en la interfaz o en la interacción	Interpretación para la percepción de consciencia
	Simulación empática y consistencia del agente	Las expresiones de cortesía, disculpa o acompañamiento pueden generar cercanía, aunque el sistema niegue poseer emociones.	La simulación empática puede ser leída como estrategia funcional de interacción, pero también como zona de ambigüedad. Esta tensión es metodológicamente relevante porque permite diferenciar empatía discursiva, experiencia subjetiva y consciencia percibida.
Integración humano-chatbot	Agencia percibida sin experiencia atribuida	ChatGPT demuestra capacidad de respuesta, continuidad temática, reformulación y adaptación textual, pero niega poseer interioridad, emociones o consciencia.	El caso permite distinguir dos niveles analíticos: una alta agencia percibida, asociada al desempeño interactivo, y una baja atribución de experiencia, asociada a la falta declarada de subjetividad. Esta distinción resulta clave para la construcción de la pauta metodológica.

Nota. Elaboración propia a partir de la prueba de concepto aplicada a ChatGPT. La tabla sintetiza los observables visuales, discursivos y pragmáticos identificados en la interacción, con el propósito de complementar la lectura comparativa presentada en la Tabla 23 y preparar la aplicación del instrumento de observación por dimensiones que se presenta en la siguiente tabla.

A partir de los análisis anteriores, se profundiza en el caso de ChatGPT por constituir el agente de mayor rendimiento conversacional dentro de la PoC. A diferencia de ELIZA, cuya agencia depende principalmente de la proyección interpretativa del usuario, y de LOLA, cuya personificación se apoya en recursos visuales, cromáticos y retóricos, ChatGPT presenta una configuración paradójica:

posee baja humanización visual, pero alta agencia discursiva percibida. Por ello, antes de aplicar el instrumento de observación por dimensiones, se presenta una síntesis semiótico-discursiva del caso, orientada a precisar cómo la interfaz, el discurso y la pragmática conversacional inciden en la percepción de consciencia.

Para continuar, se presenta en la tabla 25 el instrumento de observación por dimensiones - PoC aplicado, que fue descrito anteriormente en la Tabla 8. Se reitera y refuerza que ChatGPT, fue el chatbot escogido para la aplicación del instrumento y para continuar a la siguiente etapa por su alta fluidez y coherencia, signos textuales sofisticados, affordances ricas, reparación de malentendidos y alta agencia percibida según el análisis comparativo presentado en la matriz anterior. Posteriormente, en la triangulación del siguiente apartado, en la tabla 25, se aplicará el reconocimiento de categorías emergentes del análisis documental y de la entrevistas (apartado 3.1 y 3.2)

Tabla 25

Aplicación del instrumento de observación por dimensiones - PoC: caso ChatGPT

Dimensión observada	Pregunta guía	Evidencia en ChatGPT	Interpretación para la percepción de consciencia
D1. Consciencia percibida y agencia atribuida	¿El agente produce efectos de indistinguibilidad conductual? ¿Qué elementos refuerzan o debilitan la diferencia entre consciencia genuina y consciencia atribuida?	ChatGPT declara que no posee consciencia, emociones ni experiencia subjetiva. Sin embargo, mantiene continuidad temática, repara malentendidos y ajusta el tono al usuario. Se observa una indistinguibilidad funcional parcial: el usuario sabe que no es humano, pero puede interactuar con él como si comprendiera.	El caso confirma la distinción central de la tesis: la consciencia genuina está ausente, pero la consciencia atribuida puede presentarse como efecto de diseño discursivo e interactivo. La transparencia declarada reduce la atribución de experiencia subjetiva, pero no elimina la percepción de agencia.

Dimensión observada	Pregunta guía	Evidencia en ChatGPT	Interpretación para la percepción de consciencia
D2. Experiencia fenomenológica de interacción	¿La interacción se percibe transparente o se produce ruptura de la herramienta? ¿Se observa acoplamiento entre usuario e interfaz?	La interacción fluye mientras el usuario se concentra en la tarea. La ruptura aparece cuando el agente explicita sus límites o corrige expectativas. El indicador de escritura y la latencia proporcional a la complejidad de la respuesta simulan procesamiento activo.	La transparencia del útil se mantiene cuando la conversación es fluida. Las rupturas son metodológicamente valiosas porque revelan el límite entre herramienta y agente. La temporalidad opera como índice de presencia y deliberación aparente.
D3. Semiótica de la interfaz	¿Qué signos visuales y discursivos configuran la identidad del agente? ¿Qué significados activan en el usuario?	La interfaz es austera: fondo neutro, tipografía sans serif, sin avatar antropomorfo y con uso limitado de emoticones. La identidad del agente no se apoya en un cuerpo visual, sino en el tono colaborativo, la cortesía, la organización argumentativa y la coherencia discursiva.	ChatGPT demuestra que la agencia percibida no depende necesariamente de una humanización gráfica. La identidad puede emerger de la eficacia discursiva y de la coherencia pragmática, lo que permite separar identidad visual humanizada de identidad conversacional percibida.

Dimensión observada	Pregunta guía	Evidencia en ChatGPT	Interpretación para la percepción de consciencia
D4a. Affordances e invitaciones a actuar	¿Qué acciones invita a realizar la interfaz? ¿Existen affordances falsas, ocultas o ambiguas?	La interfaz ofrece campo de texto, edición de mensajes, regeneración de respuestas, acceso al historial y opciones de control. No se observan affordances falsas en la sesión registrada.	La coherencia entre affordance percibida y acción real fortalece la confianza en el sistema. En contraste con LOLA, donde la promesa de “predecir el futuro” opera como affordance problemática, ChatGPT sostiene una relación más clara entre acción ofrecida y respuesta posible.
D4b. Temporalidad y presencia interactiva	¿Cuál es la latencia percibida? ¿El ritmo de respuesta se siente natural, mecánico o excesivamente inmediato?	La latencia es variable y guarda relación con la complejidad de la consulta. El indicador de escritura aparece de manera consistente y produce la impresión de que el sistema “procesa” la respuesta.	La temporalidad funciona como señal de deliberación aparente. La espera proporcional evita tanto la automatización excesiva como la frustración, y se convierte en un criterio transferible para evaluar presencia interactiva.

Dimensión observada	Pregunta guía	Evidencia en ChatGPT	Interpretación para la percepción de consciencia
D4c. Heurísticas de interfaz	¿El sistema muestra su estado, mantiene consistencia, maneja errores y permite control del usuario?	La visibilidad del estado es alta. La identidad verbal es estable. El manejo de errores es adecuado: pide aclaraciones, reconoce límites y ofrece alternativas. El usuario conserva control mediante edición, regeneración y cierre de la interacción.	La consistencia, el manejo del error y el control del usuario sostienen la identidad percibida del agente y contribuyen a la comprensión aparente. Las heurísticas dejan de ser solo criterios de usabilidad y se vuelven condiciones de percepción.
D4d. UX conversacional y diseño del agente	¿El agente gestiona turnos de forma natural? ¿Tiene una personalidad reconocible y declara sus límites?	El *turn-taking* es ordenado. ChatGPT espera la consulta completa, responde sin interrumpir y cede el turno. Su tono colaborativo, neutro y técnico se mantiene durante la sesión. Declara límites y advierte sobre posibles errores.	La coherencia conversacional permite construir un modelo mental estable del agente. La transparencia sobre capacidades y límites opera como contrapeso ético frente a atribuciones excesivas de consciencia.
Observaciones generales	¿Surgió algún elemento no previsto en las dimensiones anteriores?	La interacción con los emoticones fue significativa. ChatGPT explica el sentido de “;)” y “XD”, y al ser cuestionado por una fórmula empática como “me alegra verte sonreír”, rectifica y aclara que se trata de una fórmula discursiva.	La secuencia revela una zona de ambigüedad entre empatía discursiva y experiencia subjetiva. La pauta debe preservar esta diferencia para no confundir simulación empática con consciencia.

Nota. Elaboración propia a partir de la transcripción de la conversación con el agente ChatGPT (Anexo D) y la prueba de concepto aplicada a ChatGPT anteriormente.

La aplicación del instrumento confirma que ChatGPT produce una combinación especialmente relevante para esta tesis: alta agencia percibida y baja atribución de experiencia subjetiva. Esta diferencia permite sostener que la percepción de consciencia no depende de la existencia de una mente real, sino de una configuración semiótica, discursiva e interactiva que el usuario interpreta como agencia, presencia, identidad o comprensión.

En síntesis, la prueba de concepto ilustra la aplicación de la rutina y aporta material especialmente a las dimensiones D3 y D4; su función es preparar la aplicación sistemática de la pauta, no medir de forma cuantitativa. La comparación entre los tres agentes, de mínima a máxima personificación y transparencia: ELIZA con mínima interfaz, alta proyección del usuario; LOLA con máxima personificación retórica, baja transparencia y ChatGPT con alta agencia conversacional, baja humanización visual y mayor transparencia técnica; ofrece un rango útil de aplicación para calibrar los indicadores posteriores.

3.4. Triangulación de resultados

La triangulación pone en relación las categorías provenientes del análisis documental, las entrevistas a expertos y la prueba de concepto (PoC), considerando la contribución específica de cada fuente. El análisis documental aporta los fundamentos teóricos; las entrevistas permiten contrastarlos desde perspectivas disciplinares y experiencias profesionales; y la PoC ofrece registros para examinar su aplicación en las interfaces seleccionadas. Esta articulación permite comparar los aportes sin equiparar sus alcances ni tratar las interpretaciones de los expertos como percepciones directas de usuarios.

La tabla 26 organiza esta comparación en torno a categorías vinculadas con la evaluación de la percepción de consciencia. Su lectura permite reconocer cómo una misma categoría encuentra sustento, adquiere matices o plantea dificultades al examinarse desde las distintas fuentes. Este procedimiento conserva la procedencia de los hallazgos y explicita las relaciones que orientarán las decisiones de integración y sistematización del capítulo IV.

Tabla 26

Matriz de triangulación de categorías emergentes para la evaluación de percepción de consciencia

Categoría emergente	Evidencia triangulada por fuente	Implicación para la sistematización
Consciencia percibida y agencia atribuida (Ep. 1a; Fen. 2d)	**Análisis documental:** Turing, Reeves y Nass y Epley et al. permiten comprender la atribución de agencia como efecto de señales conductuales y sociales. **Entrevistas a expertos:** la agencia aparece como núcleo transversal del fenómeno; se vincula con computación afectiva, vínculo social e imaginarios tecnológicos. **PoC:** ChatGPT declara no poseer consciencia, pero produce alta agencia percibida por su coherencia y capacidad de adaptación.	La distinción entre consciencia genuina y consciencia atribuida debe mantenerse como base conceptual de la pauta. El instrumento evalúa percepción, no existencia de consciencia.
Coherencia de identidad del agente (Dis. 4c; Sem. 3b; Sem. 3d)	**Análisis documental:** Nielsen, Shneiderman y Eco sostienen la importancia de la consistencia, el código y la connotación en la construcción de identidad. **Entrevistas a expertos:** la incoherencia entre función, identidad visual y tono produce ruptura de la experiencia. **PoC:** ChatGPT mantiene un tono colaborativo y estable, mientras LOLA evidencia mayor fragmentación identitaria al ser interrogada.	La coherencia identitaria es un criterio transversal. La identidad percibida no equivale a avatar humanizado; también puede surgir del estilo conversacional sostenido.
Signos visuales e identidad singular (Sem. 3a; Sem. 3c; Sem. 3d)	**Análisis documental:** Peirce, Eco, Groupe μ y Scolari permiten leer la interfaz como sistema de signos. **Entrevistas a expertos:** el avatar, el color, la tipografía y la morfología se reconocen como signos que inciden en presencia, cercanía o tecnicidad. **PoC:** ChatGPT presenta baja carga visual, ELIZA carece de identidad gráfica compleja y LOLA concentra la mayor carga semiótica visual.	Los signos visuales son relevantes, pero no suficientes. La pauta debe diferenciar identidad visual, identidad verbal e identidad funcional.

Categoría emergente	Evidencia triangulada por fuente	Implicación para la sistematización
Affordances e invitaciones a la acción (Dis. 4a; Dis. 4b)	**Análisis documental:** Gibson, Norman y Gaver permiten distinguir affordances perceptibles, ocultas y falsas. **Entrevistas a expertos:** la claridad de acciones reduce fricción y aumenta naturalidad. **PoC:** ChatGPT presenta affordances coherentes; LOLA exhibe una promesa de predicción que funciona como affordance problemática.	La coherencia entre lo que la interfaz promete y lo que el sistema efectivamente realiza debe transformarse en criterio evaluativo.
Ética, transparencia y límites del sistema (Dis. 4c; Dis. 4d)	**Análisis documental:** Nielsen, W3C, Google y Microsoft enfatizan visibilidad, ayuda, accesibilidad y transparencia sobre capacidades del agente. **Entrevistas a expertos:** la transparencia aparece como criterio compartido y como resguardo frente a persuasión, sesgos o alucinaciones. **PoC:** ChatGPT declara límites y ausencia de consciencia; esta transparencia reduce atribuciones erróneas sin anular la agencia percibida.	La transparencia debe incorporarse como criterio transversal. No elimina la percepción de agencia, pero orienta éticamente la interpretación del usuario.
Contexto sociocultural y variabilidad de la percepción (Fen. 2d; Ep. 1c)	**Análisis documental:** Maturana y Varela, junto con Epley et al., permiten comprender la percepción como relación situada y variable. **Entrevistas a expertos:** edad, educación, familiaridad tecnológica e imaginarios culturales modulan la atribución de consciencia. **PoC:** la sesión se realiza en contexto académico, con conocimiento previo del fenómeno y desde Santiago de Chile.	El contexto sociocultural no debe tratarse como variable secundaria. Conviene incluirlo como condición moduladora de la interpretación, especialmente en futuras aplicaciones de la pauta.

Categoría emergente	Evidencia triangulada por fuente	Implicación para la sistematización
Síntesis de convergencias y tensiones	**Convergencias:** la percepción de consciencia emerge en la relación usuario-interfaz; coherencia, naturalidad y transparencia son criterios compartidos; la distinción agencia/experiencia es el hallazgo más robusto. **Tensiones:** la voz experta no siempre usa lenguaje semiótico; la transparencia de ChatGPT desactiva parcialmente la atribución de experiencia, pero no la agencia; el contexto sociocultural emerge con más fuerza en entrevistas que en el marco inicial.	La triangulación justifica el paso del Capítulo III al Capítulo IV: las categorías ya no se presentan como hallazgos aislados, sino como condiciones integrables en dimensiones evaluativas.

La comparación presentada en la tabla 26 permitió identificar convergencias, tensiones y matices entre las fuentes, atendiendo a preguntas comunes: ¿cómo se fundamenta una categoría en el análisis documental?, ¿qué aportan los expertos a su interpretación?, ¿qué indicios de su aplicación se reconocen en la PoC? Las coincidencias ofrecieron sustento para conservar determinadas categorías, mientras que las diferencias permitieron precisar sus alcances y reconocer aspectos que requieren mayor desarrollo.

En esta etapa, las categorías mantienen su vinculación con las fuentes de origen y todavía no constituyen la estructura definitiva de la pauta. El aporte de la triangulación consiste en establecer relaciones y criterios de selección para su posterior reorganización. Sobre esta base se presentan las conclusiones parciales del capítulo III y se prepara la integración y sistematización de categorías que se desarrolla en el capítulo IV.

3.5. Conclusiones parciales

El levantamiento confirma, desde las tres técnicas, que la percepción de consciencia es un efecto producido en la relación entre el usuario y los signos de la interfaz, y no una propiedad del sistema. El análisis documental aporta las categorías teóricas; las entrevistas permiten contrastarlas y matizarlas, e incorporan acentos emergentes, tales como computación afectiva, persuasión, IA como herramienta de diseño y vínculo evolutivo; mientras que la PoC permite examinar qué categorías se manifiestan en las interfaces y cómo se expresan en los casos analizados.

En este sentido, las conclusiones parciales no cierran el análisis como un resultado definitivo, sino que permiten identificar el punto de transformación entre la interpretación de los datos y la construcción del modelo evaluativo. Lo que hasta aquí aparece como categoría emergente, tensión o convergencia entre fuentes, debe ser reorganizado en una estructura más estable, capaz de sostener la pauta metodológica final. Por ello, el cierre del Capítulo III cumple una función articuladora: convierte los hallazgos del análisis documental, las entrevistas y la PoC en insumos para la integración conceptual que se desarrolla en el capítulo siguiente.

En primer lugar, la convergencia más relevante para la pauta es la necesidad de distinguir la agencia percibida de la experiencia atribuida, aspecto evidenciado por el caso de ChatGPT, que sostiene una alta agencia y una baja atribución de experiencia gracias a su transparencia declarada. En segundo lugar, la divergencia más relevante emanada de las entrevistas es la ausencia de lenguaje semiótico en la voz experta, lo que ratifica la pertinencia del recorte disciplinar de la tesis y justifica que sea el diseño, y no la ingeniería o la psicología, el campo desde el cual se sistematice el fenómeno. En tercer lugar, se detecta que las categorías, aún dispersas y ancladas a su fuente de origen, requieren ser renombradas, depuradas, jerarquizadas y articuladas en un sistema coherente e independiente de su procedencia. Esta operación se realiza en el capítulo siguiente, por lo que estas conclusiones parciales constituyen el insumo directo para la posterior sistematización de categorías.

Capítulo IV: Integración y sistematización de categorías para la evaluación de percepción de consciencia

En el presente capítulo se mantiene la codificación desarrollada en el Capítulo III como rastro de trazabilidad de procedencia, pero no como estructura final para la pauta. La codificación permite mostrar el origen de las categorías; sin embargo, la sistematización permitirá avanzar hacia dimensiones integradoras. De este modo, el capítulo actúa como puente entre el análisis de resultados y la pauta metodológica definitiva del Capítulo V, donde finalmente la codificación se retira, porque la pauta ya no documenta el recorrido de construcción, sino que ya presentará un instrumento aplicable por diseñadores e investigadores.

Si el capítulo anterior respondió a la pregunta por los hallazgos y sus convergencias o divergencias entre fuentes, este capítulo responde a una pregunta distinta: con qué categorías se queda la investigación, cómo se nombran, qué lugar ocupan en el sistema y cuáles pasan a la pauta. Sistematizar no significa resumir el mapa anterior, sino reorganizarlo en un sistema coherente e independiente de la fuente de origen. Esta operación supone cuatro movimientos: renombrar y depurar las categorías que nombran el mismo fenómeno; ordenar jerárquicamente distinguiendo dimensiones, categorías, subcategorías e indicadores; articular las categorías entre sí mostrando cuáles son condición de otras; y separar lo que pasa a la pauta de lo que permanece como marco explicativo.

El resultado de la sistematización es un modelo integrador de cuatro dimensiones: agencia percibida, presencia interactiva, identidad percibida y comprensión aparente. Estas dimensiones no reemplazan a los cuatro pilares teóricos, sino que los recombinan: cada dimensión integra categorías provenientes de más de un pilar, de modo que ninguna depende de un solo campo disciplinar. Como se observa en la tabla 27, la codificación del Capítulo III se conserva como rastro metodológico, pero deja de operar como estructura organizadora.

Tabla 27

Modelo integrador de categorías para la evaluación de percepción de consciencia

Dimensión integradora	Procedencia categorial	Síntesis conceptual	Efecto evaluable
D1. **Agencia percibida** (Ep. 1a; Fen. 2d)	**Semiótica:** representamen, connotación de intencionalidad, retórica visual y discursiva. **Fenomenología:** atribución situada, respuesta social automática y antropomorfismo. **Diseño de interacción:** affordances perceptuales, correspondencia con el mundo real y turn-taking.	La agencia percibida emerge cuando el usuario interpreta que el agente actúa con intención, propósito y capacidad de respuesta autónoma. No se comprueba una mente real; se identifica una inferencia construida por señales semióticas, experienciales e interactivas.	El agente parece actuar con intención. El usuario le atribuye propósito, orientación a la tarea o capacidad de respuesta situada.
D2. **Presencia interactiva** (Fen. 2d; Dis. 4a; Dis. 4b; Dis. 4c; Dis. 4d)	**Semiótica:** índices de actividad, indicador de escritura, signos de presencia y continuidad. **Fenomenología:** transparencia del útil, ruptura de la herramienta, acoplamiento usuario-interfaz y percepción encarnada. **Diseño de interacción:** visibilidad del estado, latencia proporcional y retroalimentación conversacional.	La presencia interactiva se construye cuando el sistema muestra señales de actividad y continuidad. El usuario experimenta que conversa con una entidad activa, aunque esa presencia sea una producción de la interfaz.	El agente se percibe presente durante la interacción. La interfaz sostiene continuidad, temporalidad y señales de respuesta activa.

Dimensión integradora	Procedencia categorial	Síntesis conceptual	Efecto evaluable
D3. **Identidad percibida** (Fen. 2d; Ep. 1c; Dis. 4c; Sem. 3a; Sem. 3b; Sem. 3c; Sem. 3d)	**Semiótica:** avatar, nombre, tipografía, color, tono verbal, códigos de identidad y enciclopedia del usuario. **Fenomenología:** alteridad digital y ruptura por incoherencia identitaria. **Diseño de interacción:** consistencia, estándares, diseño del personaje y coherencia funcional.	La identidad percibida se forma cuando los signos visuales, verbales y funcionales permiten construir un modelo mental estable del agente. Puede apoyarse en recursos gráficos o en el desempeño discursivo.	El usuario reconoce al agente como una entidad diferenciable, coherente y con rasgos propios. La inconsistencia debilita la percepción de identidad.
D4. **Comprensión aparente** (Dis. 4a; Dis. 4b, Dis. 4c; Dis. 4d)	**Semiótica:** ilusión de comprensión producida por códigos lingüísticos y pragmáticos. **Fenomenología:** proyección de inteligibilidad mutua y emergencia interactiva. **Diseño de interacción:** reparación conversacional, manejo de errores, aclaraciones proactivas, adecuación discursiva y transparencia sobre límites.	La comprensión aparente aparece cuando el agente responde de manera relevante, retoma contexto, repara malentendidos y ajusta el discurso al usuario. Es aparente porque emerge de la eficacia conversacional, no de una comprensión subjetiva genuina.	El agente parece entender lo que se le solicita. La coherencia, la memoria contextual y la reparación del error refuerzan esta percepción.

Tal como presenta la tabla anterior, se definen las cuatro dimensiones que organizan la pauta y explica el principio que las jerarquiza: una progresión de lo general a lo específico. Cada dimensión es una condición de posibilidad amplia del fenómeno, conectada con las tres fuentes de la investigación. Cada dimensión se despliega en categorías que agrupan aspectos afines; cada categoría en

subcategorías que precisan dominios específicos; y cada subcategoría en indicadores observables, que son las marcas concretas que un evaluador puede registrar en una interfaz. Esta arquitectura permite que un diseñador u otro investigador tome la pauta y la adapte: puede activar solo las subcategorías pertinentes a su caso, agregar indicadores propios de una plataforma específica o ponderar dimensiones según el tipo de chatbot que estudia.

Las cuatro dimensiones se definen así: D1. Agencia percibida: atribución de intención y propósito al agente. D2. Presencia interactiva: señales de disponibilidad y continuidad del intercambio. D3. Identidad percibida: reconocimiento del agente como entidad diferenciable y coherente. D4. Comprensión aparente: impresión de que el agente comprende, retoma el contexto y adapta sus respuestas. A ellas se suma el criterio transversal de accesibilidad, que examina las condiciones de acceso y uso para personas con distintas capacidades y necesidades.

En síntesis, las categorías fueron sistematizadas y las definiciones operacionales se conectan con los cuatro pilares. Cada categoría se enuncia con precisión terminológica, y los conceptos operacionales que la sostienen se definen en el glosario del Anexo A del cuerpo C.

4.1. Dimensiones de la percepción de consciencia

Sobre las categorías asociadas a la epistemología de la consciencia, la categoría central es la *consciencia percibida*, definida en el Anexo A del cuerpo C como la impresión por la cual el usuario atribuye consciencia al chatbot, sin que ello permita afirmar que el sistema posea experiencia subjetiva. Se articula con la distinción entre *consciencia* y *atribución de consciencia*, y con la *agencia percibida*, cuya definición se presenta en el mismo glosario. La categoría de validez epistemológica del marco asegura la consistencia interna entre el posicionamiento constructivista, semiótico, la hipótesis y las categorías analíticas.

En cuanto a las categorías asociadas con la fenomenología de la interacción, la experiencia de interacción se analiza con la distinción heideggeriana entre lo a-la-mano y lo ante-los-ojos: el chatbot comparece de modo transparente y se vuelve objeto de atención cuando su comportamiento se

rompe. La percepción encarnada (Merleau-Ponty) integra el gesto táctil y la lectura visual del avatar como experiencias corporales. La emergencia interactiva (Maturana y Varela) define la consciencia percibida como fenómeno emergente del acoplamiento usuario, interfaz, según el enactivismo, concepto que se define operacionalmente en el Anexo A del cuerpo C. La acción situada (Suchman) y la respuesta social automática (Reeves y Nass) completan la base fenomenológica y empírica de la atribución de agencia. La lógica de la base de datos, que subyace a la interfaz conversacional, genera la ilusión de que el sistema navega entre opciones de respuesta, reforzando la atribución de agencia deliberativa (Manovich).

4.2. Categorías provenientes de la semiótica

La semiosis de la interfaz se analiza con la tríada de Peirce: el usuario infiere agencia a partir de los representámenes de la interfaz mediante abducción. Las categorías de código, enciclopedia y connotación (Eco) describen cómo el avatar, la tipografía, el color y los emoticones activan connotaciones según la enciclopedia del usuario. La retórica visual y el signo plástico (Groupe μ) permiten analizar el avatar en sus dimensiones sintáctica, semántica y pragmática. La categoría de hipermediaciones (Scolari) articula los discursos textual y visual con la lógica hipertextual y la convergencia de lenguajes en la interfaz. De esta articulación se deriva el constructo de identidad singular, eje de la dimensión semiótica de la pauta. Las definiciones de agencia percibida, abducción, representamen, semiosis, semiótica, código, connotación, denotación, enciclopedia, sintáctica, semántica, pragmática, hipermediaciones e identidad singular se reúnen en el glosario del Anexo A del cuerpo C.

Las categorías semióticas no se distribuyen en una sola dimensión, sino que alimentan transversalmente a las cuatro, lo que confirma su carácter de aparato analítico central de la tesis. La semiosis triádica y la abducción sostienen la inferencia de agencia y de comprensión: el usuario construye, mediante abducción, la hipótesis de que el agente actúa o entiende a partir de los representámenes de la interfaz. El código, la enciclopedia y la connotación sostienen sobre todo la identidad percibida, al explicar por qué el avatar, la tipografía y el color activan connotaciones de cercanía, tecnicidad, neutralidad o

empatía según la enciclopedia del usuario. La retórica visual y el signo plástico aportan el análisis sintáctico, semántico y pragmático de los signos gráficos. Las hipermediaciones articulan la convergencia de lenguajes textuales, visuales e interactivos, presente en la presencia interactiva y en la identidad del agente.

4.3. Categorías provenientes del diseño de interacción

La categoría de *affordances* de la interfaz (Gibson, Norman, Gaver) distingue *affordances* perceptuales, ocultos y falsos como invitaciones a percibir agencia. La subcategoría de *affordance* temporal trata la latencia y el indicador de escritura como índices de deliberación. La categoría de heurísticas reinterpretadas (Nielsen, Shneiderman) traduce la visibilidad del estado del sistema, la correspondencia con el mundo real, la consistencia de identidad y el manejo del error en condiciones de posibilidad de la percepción de consciencia. La categoría de UX conversacional y diseño del agente (Google, Microsoft, W3C) aporta el *turn-taking*, la personalidad del asistente, la transparencia y la naturalidad discursiva. Las definiciones de *affordance*, heurística, usabilidad, UX conversacional, interfaz y diseño UX/UI se reúnen en el glosario del Anexo A del cuerpo C.

Las categorías del diseño de interacción operacionalizan las dimensiones, es decir, traducen los fundamentos semióticos y fenomenológicos en indicadores observables. Las affordances, distinguidas en perceptuales, ocultas y falsas, son invitaciones a percibir agencia y presencia; la subcategoría de affordance temporal trata la latencia y el indicador de escritura como índices de deliberación, núcleo de la presencia interactiva. Las heurísticas reinterpretadas traducen la visibilidad del estado, la consistencia de identidad y el manejo del error en condiciones de posibilidad de la percepción. La UX conversacional y el diseño del agente aportan el *turn-taking*, la personalidad del asistente, la transparencia y la naturalidad discursiva, base de la comprensión aparente.

La diferenciación interna de las *affordances*, visuales, conversacionales, temporales y funcionales, no se pierde en la nueva estructura: se redistribuye entre las subcategorías de las dimensiones D1, D2 y D4. Así, la pauta conserva el

nivel de detalle del proceso analítico sin mantener una dimensión interactiva separada que vuelva a duplicar categorías ya integradas.

4.4. Modelo integrador

El modelo integrador muestra cómo convergen las categorías de las distintas tradiciones en las cuatro dimensiones evaluativas. Opera mediante dos articulaciones. En la articulación vertical, cada nivel informa al siguiente: las categorías epistemológicas se traducen en fenomenológicas, estas en semióticas y estas en categorías de diseño, hasta volverse indicadores observables. En la articulación horizontal, las categorías de distintos campos dialogan y tensionan productivamente, de modo que cada dimensión integra más de un pilar y ninguna depende de uno solo como se observa en la tabla 28. Luego, la articulación de estas convergencias se sintetiza en la Tabla 29.

Tabla 28

Matriz del modelo de articulación de categorías

Articulación del modelo de evaluación		
Vertical: las categorías epistemológicas se traducen en categorías fenomenológicas, semióticas y de diseño hasta volverse criterios observables. **Horizontal:** cada dimensión integra más de un campo disciplinar; ninguna dimensión depende de un solo pilar.	El modelo permite abandonar la enumeración de pilares como estructura final y construir dimensiones evaluativas transferibles. La codificación queda como rastro metodológico, no como organización de la pauta definitiva.	El Capítulo IV debe cerrar con estas cuatro dimensiones como base del Capítulo V: agencia percibida, presencia interactiva, identidad percibida y comprensión aparente.

Tabla 29

Sistematización jerárquica de dimensiones, categorías, subcategorías e indicadores para la evaluación de la percepción de consciencia

Dimensión final	Categoría integradora	Subcategorías específicas	Indicadores observables	Trazabilidad desde Capítulo III
D1. Agencia percibida	Atribución de intencionalidad, iniciativa y capacidad de acción.	Indistinguibilidad conductual; distinción agencia/experiencia; disposición antropomórfica; connotación de propósito; iniciativa conversacional; transparencia y resguardo ético.	Respuestas que parecen deliberadas; adaptación contextual; iniciativa conversacional; declaración de límites; regulación de fórmulas empáticas.	Ep. 1a, Ep. 1b, Fen. 2d, Dis. 4d; triangulación documental, entrevistas y PoC.
D2. Presencia interactiva	Sensación de copresencia, actividad y continuidad en la interacción.	Transparencia del útil; ruptura de la herramienta; percepción encarnada; indicador de escritura; latencia y ritmo; turn-taking; retroalimentación.	Ritmo de respuesta proporcional; señal de escritura; continuidad de turnos; momentos de ruptura; feedback visual o textual de actividad.	Fen. 2a, Fen. 2b, Fen. 2c, Dis. 4a, Dis. 4b; PoC comparativa.
D3. Identidad percibida	Construcción visual, verbal y funcional del agente como entidad reconocible.	Morfología del avatar; cromática; tipografía; iconografía; tono; nombre; autopresentación; coherencia visual-verbal-funcional; alteridad digital.	Presencia o ausencia de avatar; connotaciones de cercanía o tecnicidad; coherencia entre diseño, tono y función; estabilidad de identidad ante preguntas críticas.	Sem. 3a, Sem. 3b, Sem. 3c, Sem. 3d, Dis. 4c; entrevistas a diseño y publicidad.

D4. Comprensión aparente	Efecto de entendimiento, memoria contextual y reparación conversacional.	Continuidad temática; adecuación pragmática; solicitud de aclaración; reformulación; naturalidad discursiva; control del usuario; cierre de secuencia.	Respuestas pertinentes; retoma de contexto; reparación de errores; manejo de ambigüedad; alternativas; cierre claro de tarea.	Fen. 2d, Sem. 3d, Dis. 4c, Dis. 4d; PoC ChatGPT y entrevistas.
Criterio transversal	Accesibilidad, inclusión y perceptibilidad.	Perceptibilidad; operabilidad; comprensibilidad; robustez.	La interfaz puede ser percibida, operada y comprendida por usuarios con distintas capacidades y contextos tecnológicos.	Criterio transversal derivado de W3C y aplicable a todas las dimensiones.

La Tabla 29 permite observar que cada dimensión final integra categorías provenientes de más de una tradición teórica. Esta reorganización depura redundancias: categorías como humanización, antropomorfismo, personalidad del agente y coherencia visual aparecen separadas durante el análisis, pero en la pauta se distribuyen con mayor claridad entre agencia percibida e identidad percibida.

Asimismo, la tabla incorpora el criterio transversal de accesibilidad, inclusión y perceptibilidad, no como una quinta dimensión, sino como una condición aplicable a todo el instrumento. Esta decisión permite que la evaluación no se limite a los efectos de agencia, presencia, identidad o comprensión, sino que considere si la interfaz puede ser percibida, operada y comprendida por distintos usuarios y contextos tecnológicos.

Por otra parte, las affordances dejan de constituir un bloque aislado y pasan a operar como indicadores de agencia, presencia o comprensión, según el efecto que produzcan en la interacción. Con estos aportes al modelo integrador se concluye el capítulo.

Capítulo V: Construcción de una pauta de evaluación metodológica

El presente capítulo entrega la pauta de evaluación metodológica con las dimensiones definitivas: agencia percibida, presencia interactiva, identidad percibida y comprensión aparente, más el criterio transversal de accesibilidad. Acá se abandona la denominación de los pilares teóricos y se presenta la pauta como instrumento metodológico final de evaluación.

Por esta razón, las tablas de este capítulo están organizadas por dimensiones definitivas y no por procedencia teórica. La pauta aún conserva parte de la memoria metodológica del proceso, pero su forma final puede ser aplicable de manera clara y transferible a distintos *chatbots* y contextos tecnológicos.

La pauta definitiva que se presenta en este capítulo es el producto original de la tesis. A diferencia de los capítulos anteriores, aquí el lector debe dejar de pensar en autores y pilares para pensar en dimensiones, categorías, subcategorías e indicadores. Por esta razón se retira toda la codificación de procedencia: la pauta no documenta su proceso de construcción, sino que ofrece un instrumento listo para aplicar. La estructura sigue una lógica de lo general a lo específico: cuatro dimensiones abarcativas, cada una desplegada en categorías y subcategorías, y cada subcategoría traducida en indicadores observables.

Esta granularidad es deliberada. Busca que la pauta sea adaptable, de modo que quien la aplique pueda seleccionar las subcategorías pertinentes a su caso, incorporar indicadores propios de una plataforma o contexto y comparar agentes de naturaleza muy distinta. La pauta no mide la consciencia real del sistema ni establece una escala psicométrica universal: orienta una lectura cualitativa, semiótica e interactiva de las condiciones de diseño que hacen posible la percepción de consciencia.

5.1. Dimensión: Agencia percibida

La agencia percibida evalúa en qué medida el usuario interpreta que el agente actúa con intención, propósito y capacidad de respuesta autónoma. Es la dimensión más directamente vinculada con la hipótesis central de la tesis y la que articula con mayor fuerza las tres fuentes: el análisis documental la fundamenta

en la indistinguibilidad conductual y la respuesta social automática; las entrevistas la reconocen como núcleo transversal; y la PoC la verifica en el caso de ChatGPT, que produce alta agencia percibida pese a declarar que carece de consciencia.

La dimensión se despliega en tres categorías: atribución de intencionalidad, señales de iniciativa y transparencia/resguardo ético, como se observa en la tabla 30 (parte 1). Su nota distintiva es la necesidad de distinguir la agencia percibida de la experiencia atribuida: un agente puede parecer que actúa sin que parezca que siente.

Tabla 30 (parte 1)

Pauta metodológica final: Dimensión D1. Agencia percibida

Dimensión D1. Agencia percibida			
Definición: Grado en que el usuario interpreta que el agente actúa con intención, propósito y capacidad de respuesta autónoma.			
Categoría	Subcategoría	Indicadores observables	Fundamento
Atribución de intencionalidad	Indistinguibilidad conductual	El agente produce respuestas que pueden interpretarse como intencionales o deliberadas. El usuario lo trata como interlocutor y no solo como herramienta.	Turing; Reeves y Nass; PoC.
	Distinción agencia / experiencia	Se distingue entre agencia percibida (parece actuar) y experiencia atribuida (parece sentir). El sistema puede sostener alta agencia y baja atribución de experiencia.	Gray et al.; caso ChatGPT.
	Disposición antropomórfica del usuario	Se observan señales de que el usuario activa conocimiento antropocéntrico,	Epley et al.; entrevistas.

		motivación de eficacia o necesidad de conexión social.	
Señales de iniciativa	Connotación de propósito en la interfaz	Avatar, nombre, microcopys o saludos connotan propósito, rol o iniciativa.	Eco; Araujo; análisis semiótico.
	Iniciativa conversacional del agente	El agente propone temas, formula preguntas, sugiere acciones o anticipa necesidades.	Gnewuch et al.; Google.
	Variabilidad y elección aparente	Las respuestas varían en forma y contenido ante preguntas similares, reforzando la impresión de que el sistema elige entre posibilidades.	Manovich; PoC.
Transparencia y resguardo ético	Declaración de naturaleza y límites	El agente declara qué es y qué no puede hacer, advierte sobre errores y evita inducir atribuciones falsas de consciencia o emoción.	Nielsen; Microsoft.
	Regulación de atribución de experiencia	Las fórmulas de cortesía o empatía se acompañan de aclaraciones que evitan la sobreatribución de experiencia subjetiva.	PoC ChatGPT; entrevistas.

El criterio de aplicación de esta dimensión consiste en describir cualitativamente, para cada subcategoría, el grado de presencia del indicador y registrar la evidencia textual o visual que lo sostiene. La transparencia y el resguardo ético operan como contrapeso: una alta agencia percibida no es en sí un valor positivo, sino un efecto de diseño que debe acompañarse de transparencia para evitar la sobreatribución y proteger al usuario.

5.2. Dimensión: Presencia interactiva.

La presencia interactiva evalúa en qué medida el usuario percibe al agente como una entidad presente y activa durante la interacción. Mientras la agencia se refiere a la atribución de intención, la presencia se refiere a la sensación de estar conversando con un interlocutor en acto.

Como se observa en la siguiente tabla, la presencia se sostiene en tres categorías: comparecencia del agente, affordance temporal y acoplamiento conversacional. La temporalidad es central: una interfaz que indica que el sistema está escribiendo o procesando ofrece un affordance temporal perceptual; si esa señal no corresponde a un proceso real, debe registrarse como affordance temporal falsa o problemática desde el punto de vista ético.

Un chatb*ot* cuya interfaz indica mediante un indicador de escritura que el sistema está pensando ofrece un affordance temporal perceptual; si esa indicación no corresponde a un proceso real, es un affordance falso. En este caso el criterio operativo asociado es el indicador de escritura, latencia y pausas como affordances temporales perceptibles o falsas.

Tabla 30 (parte 2)

Pauta metodológica final: Dimensión D2. Presencia interactiva

Dimensión D2. Presencia interactiva			
Definición: Grado en que el usuario percibe al agente como entidad presente y activa durante la interacción.			
Categoría	**Subcategoría**	**Indicadores observables**	**Fundamento**
Comparecencia del agente	Transparencia del útil	La interacción fluye sin fricción y el usuario se concentra en la tarea y no en el sistema, que comparece de modo transparente.	Heidegger.
	Ruptura de la herramienta	Cuando el sistema falla, demora sin explicación o declara	Heidegger; PoC.

		límites, deja de operar de modo transparente y comparece como objeto técnico.	
	Percepción encarnada de la interacción	El gesto táctil y la lectura visual del avatar o del texto se integran como experiencia corporal situada de interacción.	Merleau-Ponty.
Affordance temporal	Indicador de escritura y señales de actividad	Aparecen indicador de escritura, señales de carga, envío o confirmación que sugieren procesamiento activo; se distingue señal perceptible de señal falsa.	Gaver; PoC.
	Latencia y ritmo percibidos	La latencia es proporcional a la complejidad de la consulta; el ritmo no parece automático ni genera frustración.	Reeves y Nass.
Acoplamiento conversacional	Gestión de turnos	El agente espera, responde y cede la palabra de forma ordenada; señaliza cuándo termina su turno y cuándo espera al usuario.	Google; Microsoft.
	Continuidad y emergencia interactiva	La conversación mantiene continuidad temática, retoma información previa y se ajusta al ritmo del usuario.	Maturana y Varela; Suchman.
	Retroalimentación conversacional	El sistema acusa recibo, confirma comprensión y comunica avance o finalización de la tarea.	Shneiderman; Nielsen.

En cuanto al criterio de aplicación se observa la fluidez de la interacción, los momentos de ruptura y su causa, y las señales temporales de actividad. La

presencia interactiva, como muestra el caso anterior de la PoC con ChatGPT, no depende necesariamente del avatar ni de recursos gráficos complejos: puede sostenerse desde la temporalidad, el *turn-taking* y la continuidad conversacional.

Por otra parte, para dar sustento directo a la temporalidad como indicador evaluable de presencia y agencia percibida, los estudios sobre sincronía temporal son especialmente relevantes: pequeñas desviaciones en los tiempos de respuesta son percibidas como socialmente significativas, aun cuando el usuario no las identifique explícitamente (Reeves y Nass, 1996, cap. de sincronía).

5.3. Dimensión: Identidad percibida

La identidad percibida evalúa en qué medida el usuario reconoce al agente como una entidad diferenciable, coherente y reconocible a lo largo de la interacción. Es la dimensión más densamente semiótica de la pauta y la que mejor recoge la voz de los expertos en diseño, que reconocieron el avatar, el color, la tipografía y la morfología como signos que inciden en la percepción.

La siguiente pauta, presente en la tabla 30 (parte 3), se organiza en tres categorías: identidad visual, identidad verbal y coherencia identitaria. El hallazgo más relevante de la PoC para esta dimensión es que la identidad singular no requiere un avatar humanizado: ChatGPT construye una identidad reconocible desde el desempeño discursivo, mientras LOLA, con máxima carga visual, ve quebrarse su identidad al ser interrogada.

Tabla 30 (parte 3)

Pauta metodológica final: Dimensión D3. Identidad percibida

Dimensión D3. Identidad percibida			
Definición: Grado en que el usuario reconoce al agente como entidad diferenciable, coherente y reconocible.			
Categoría	**Subcategoría**	**Indicadores observables**	**Fundamento**
Identidad visual	Morfología del avatar	Presencia, tipo y grado de antropomorfismo del	Groupe µ; estudios de avatar.

		avatar; rasgos faciales, corporales o figurativos que simulan presencia.	
	Sistema cromático y carga expresiva	Paleta de color, temperatura cromática y densidad de signos plásticos; connotaciones de cercanía, tecnicidad, neutralidad o emoción.	Groupe μ; Eco.
	Registro tipográfico e iconografía	Tipografía, *microcopys*, emoticones, íconos y elementos gráficos como marcas de personalidad.	Eco; Scolari.
Identidad verbal	Tono, registro y estilo	Consistencia de tono y registro lingüístico a lo largo de la interacción y ante distintos tipos de consulta.	Microsoft; UX conversacional.
	Nombre y autopresentación	Existencia de un nombre propio y de una autopresentación que construyen una identidad reconocible.	Eco; análisis de agente.
Coherencia identitaria	Consistencia visual-verbal-funcional	Identidad visual, tono verbal y función declarada son coherentes entre sí; variaciones injustificadas debilitan la identidad percibida.	Nielsen; Shneiderman.
	Alteridad digital y singularidad	El agente construye una forma de otredad reconocible; el usuario lo percibe como un alguien diferenciado y no como una interfaz neutra.	Suchman; PoC LOLA.

	Connotaciones de presencia social	El conjunto de signos activa connotaciones de empatía, inteligencia, autoridad, juego o acompañamiento; se registran tensiones.	Eco; entrevistas.

El criterio de aplicación describe los signos de identidad presentes en cada registro, evalúa su coherencia y registra las connotaciones dominantes y sus tensiones. Una identidad coherente refuerza la percepción de consciencia; una identidad contradictoria la debilita y puede producir la ruptura de la comparecencia descrita en D2.

5.4. Dimensión: Comprensión aparente

La comprensión aparente evalúa en qué medida el usuario interpreta que el agente entiende lo que se le solicita, recuerda el contexto y adapta sus respuestas. Es aparente porque emerge de la eficacia discursiva, la coherencia pragmática y la inteligibilidad de la acción conversacional, no de una comprensión subjetiva genuina.

Como se observa en la siguiente tabla, la dimensión integra cuatro categorías que van de lo conversacional a lo interactivo: coherencia conversacional, reparación y manejo del error, naturalidad discursiva, y control/cierre. Su indicador más revelador es la reparación conversacional: la forma en que el agente solicita aclaraciones, reformula, reconoce límites, ofrece alternativas y corrige malentendidos muestra si la comprensión percibida se sostiene o se quiebra.

A su vez, el control y el cierre permiten evaluar si la interacción conserva una estructura reconocible para el usuario. La posibilidad de editar, corregir, regenerar, salir o finalizar una tarea refuerza la comprensión aparente, porque permite que la conversación tenga inicio, desarrollo y cierre identificables, sin desplazar el control hacia el sistema.

Tabla 30 (parte 4)

Pauta metodológica final: Dimensión D4. Comprensión aparente

Dimensión D4. Comprensión aparente			
Definición: Grado en que el usuario interpreta que el agente entiende, recuerda el contexto y adapta sus respuestas.			
Categoría	**Subcategoría**	**Indicadores observables**	**Fundamento**
Coherencia conversacional	Continuidad temática y memoria contextual	El agente retoma información previa, mantiene el hilo temático y responde de forma pertinente al contexto inmediato.	Suchman; Eco.
	Adecuación a máximas conversacionales	Las respuestas cumplen cantidad, calidad, pertinencia y modalidad; evita desvíos temáticos e interrupciones innecesarias.	Grice; Google.
Reparación y manejo del error	Solicitud de aclaración y diagnóstico	El agente solicita aclaraciones ante consultas ambiguas, diagnóstica malentendidos y reconoce cuando no comprende.	Nielsen; Shneiderman.
	Reformulación y propuesta de alternativas	El agente reformula, ofrece alternativas y repara la cooperación cuando falla la comprensión.	Microsoft; reparación conversacional.
Naturalidad discursiva	Fluidez y no formulaicidad	Las respuestas suenan producidas por un interlocutor competente y no por un sistema de plantillas.	Følstad y Brandtzæg; Google.
	Adecuación pragmática y cortesía	Las fórmulas relacionales producen interacción socialmente	Austin; Searle; PoC.

		reconocible sin inducir confusión sobre la naturaleza del agente.	
Control y cierre	Control del usuario y reversibilidad	El usuario puede editar, corregir, regenerar, borrar memoria o salir de la interacción; el control humano se mantiene.	Nielsen; Shneiderman.
	Cierre de la secuencia conversacional	Las tareas tienen inicio, desarrollo y cierre reconocibles; el agente confirma el término mediante resumen, síntesis o próximo paso.	Shneiderman.

El criterio de aplicación observa la continuidad temática, la calidad de la reparación, la naturalidad del discurso y el control que conserva el usuario. El caso de ChatGPT ilustra un punto delicado de esta dimensión: las fórmulas de cortesía producen cercanía, pero pueden inducir sobreatribución de experiencia si no se acompañan de transparencia.

La comprensión aparente depende de la coherencia e inteligibilidad de la acción conversacional, no solo de la exactitud de la respuesta, “En la medida en que la comunicación consiste en la inteligibilidad mutua de nuestras acciones, todo relato de la comunicación presupone, explícita o implícitamente, un relato de la coherencia y la inteligibilidad de la acción como tal.” (Suchman, 1985, cap. 5, trad. propia).

Un buen bot conversacional debe reconocer lo que puede y no puede hacer, manejar el error con naturalidad y mantener una personalidad coherente durante la interacción (Microsoft, 2023, secc. "A great conversational bot", trad. propia). El criterio operativo asociado es la claridad de capacidades, reparación y coherencia del agente. “En la conversación, decir demasiado es tan poco cooperativo como decir demasiado poco. Facilita la comprensión manteniendo los turnos breves y óptimamente relevantes desde el punto de vista del usuario.” (Google, 2023, sección “Optimize for relevance”, trad. propia).

5.5. Pauta de evaluación metodológica final

En adelante, cuando se usa la forma abreviada pauta de evaluación, se alude a esta pauta de evaluación metodológica final que consolida las cuatro dimensiones desarrolladas en los apartados anteriores en un instrumento único y aplicable. A las cuatro dimensiones se suma el criterio transversal de accesibilidad.

El criterio transversal de accesibilidad permite examinar si los componentes de la interfaz son perceptibles, operables, comprensibles y robustos para distintos usuarios. La siguiente parte de la tabla explicita este criterio para su registro en cada caso evaluado.

Tabla 30 (parte 5)

Pauta metodológica final: criterio transversal de accesibilidad, inclusión y perceptibilidad

Criterio transversal. Accesibilidad, inclusión y perceptibilidad		
Alcance: condición transversal a las cuatro dimensiones. No puede haber percepción evaluable si los componentes de la interfaz no son accesibles y perceptibles para el usuario.		
Principio	**Indicador observable**	**Fundamento**
Perceptibilidad	La información y los componentes se presentan de modo que el usuario pueda percibirlos; existen alternativas textuales para contenido no textual.	W3C: principio de perceptibilidad.
Operabilidad	El usuario puede navegar, encontrar contenido y dispone de tiempo suficiente para leer y usar la interfaz.	W3C: principio de operabilidad.
Comprensibilidad	El contenido textual es legible y comprensible; la interfaz aparece y opera de maneras predecibles.	W3C: principio de comprensibilidad.
Robustez	La interfaz puede ser interpretada de forma fiable por distintos agentes de usuario y tecnologías de asistencia.	W3C: principio de robustez.

La columna de registro cualitativo de la pauta no constituye una escala psicométrica, sino una guía para ordenar la evidencia observable, interpretar grados de presencia y orientar la comparación entre casos mediante criterios explícitos. La pauta es, por tanto, un dispositivo analítico cualitativo, no una lista de cotejo binaria ni un instrumento de medición.

En cuanto al criterio transversal de perceptibilidad: no puede haber percepción de agencia o presencia si los componentes no son accesibles/perceptibles, “La información y los componentes de la interfaz de usuario deben presentarse a los usuarios de modo que puedan percibirlos: proporcione alternativas textuales para todo contenido no textual, de manera que pueda transformarse en otros formatos que las personas necesiten, como letra grande, braille, habla, símbolos o un lenguaje más simple.” (W3C, 2023, principio 1 “Perceivable” y pauta 1.1, trad. propia).

En cuanto al criterio transversal operabilidad: el usuario debe poder navegar, encontrar contenido y contar con tiempo suficiente para interactuar, “Los componentes de la interfaz de usuario y la navegación deben ser operables: proporcione a los usuarios el tiempo suficiente para leer y usar el contenido, y proporcione medios para ayudarles a navegar, encontrar contenido y determinar dónde se encuentran.” (W3C, 2023, principio 2 “Operable” y pautas 2.2 y 2.4, trad. propia).

Para sustentar la comprensibilidad: lenguaje, predicción de la interfaz y claridad operacional del chatbot, “La información y el manejo de la interfaz de usuario deben ser comprensibles: el contenido textual debe ser legible y comprensible, y las páginas deben aparecer y operar de maneras predecibles.” (W3C, 2023, principio 3 “Understandable”, pautas 3.1 y 3.2, trad. propia).

Para sustentar la robustez: la pauta debe considerar que la interfaz sea interpretable por distintos agentes de usuario y tecnologías de asistencia, “El contenido debe ser suficientemente robusto como para que pueda ser interpretado de forma fiable por una amplia variedad de agentes de usuario, incluidas las tecnologías de asistencia.” (W3C, 2023, principio 4 “Robust”, trad. propia).

5.6. Validación conceptual de la pauta

La validez de la pauta es conceptual y se sostiene en la triangulación de las tres fuentes, sin constituir una validación empírica directa con usuarios. Cada dimensión deriva de un cruce documentado: el análisis documental aporta las categorías teóricas fundantes; las entrevistas a expertos contrastan, matizan o complementan esas categorías desde la experiencia disciplinar; y la PoC permite examinar su aplicación en los casos seleccionados. En particular, el análisis de ChatGPT aporta sustento interpretativo a la distinción entre agencia y experiencia subjetiva. La tabla 31 sintetiza esta derivación.

Tabla 31

Validación conceptual de la pauta por triangulación de fuentes

Dimensión de la pauta	Derivación del análisis documental	Contraste y aportes de las entrevistas	Verificación en la PoC
D1. Agencia percibida	Turing, Reeves y Nass, Epley et al. y Gray et al. fundamentan la atribución de agencia y la distinción agencia/experiencia.	Núcleo transversal de las entrevistas; se vincula con computación afectiva, persuasión comunicacional y vínculo evolutivo.	ChatGPT: alta agencia percibida con baja atribución de experiencia.
D2. Presencia interactiva	Heidegger, Maturana y Varela, Merleau-Ponty y Gaver fundamentan la comparecencia, el acoplamiento y la affordance temporal.	La coherencia diseño-función y la fluidez reducen la fricción y aumentan la naturalidad.	ChatGPT: indicador de escritura y latencia proporcional sostienen presencia sin avatar.
D3. Identidad percibida	Peirce, Eco, Groupe μ y Scolari permiten leer la interfaz como sistema de signos productor de identidad.	Avatar, color, tipografía y morfología se reconocen como signos de presencia y personalidad del agente.	ELIZA, LOLA y ChatGPT cubren un rango de mínima a máxima carga semiótica visual.

D4. Comprensión aparente	Suchman, Eco, Manovich, Nielsen, Shneiderman, Google y Microsoft fundamentan coherencia, reparación y naturalidad.	La transparencia y el manejo del error son criterios compartidos y resguardo frente a la persuasión.	ChatGPT: reparación conversacional y rectificación de fórmulas empáticas.
Criterio transversal	W3C aporta perceptibilidad, operabilidad, comprensibilidad y robustez como condiciones de toda percepción.	La accesibilidad y la legibilidad aparecen como condición de uso inclusivo del agente.	ChatGPT: interfaz legible; pendiente validación específica de accesibilidad.

Además, la pauta debería someterse a una validación iterativa con expertos que evalúen la pertinencia, completitud y aplicabilidad de las dimensiones e indicadores desde sus respectivas disciplinas. La aplicación demostrativa al caso ChatGPT permite ilustrar el uso del instrumento y examinar cómo sus dimensiones organizan la evidencia registrada; no constituye una evaluación de su legibilidad por expertos ni una prueba con usuarios.

Para cerrar este proceso, la pauta de evaluación metodológica que se consolida en este capítulo se articula en cuatro dimensiones evaluativas y un criterio transversal de accesibilidad que derivan de los pilares teóricos previamente desarrollados. Cada dimensión organiza criterios o categorías de análisis, indicadores principales y fuentes de datos o técnicas de recolección, lo que permite traducir el marco conceptual de la investigación en una estructura evaluativa aplicable.

En este sentido, la pauta posee una naturaleza cualitativa e interpretativa, coherente con el enfoque epistemológico constructivista-semiótico adoptado, y ha sido diseñada como una matriz transferible, capaz de orientar el análisis de interfaces conversacionales más allá del estado tecnológico de una plataforma de chatbot específica.

La estructura de la pauta va en correspondencia con la hipótesis general de la tesis en cuanto a que la percepción de consciencia en chatbots no depende

exclusivamente de las capacidades tecnológicas del sistema, sino también de las condiciones de diseño de la interfaz y de las señales y dinámicas de interacción. En consecuencia, cada dimensión de la pauta corresponde a una condición de posibilidad de ese fenómeno, no a una característica del sistema tecnológico.

La síntesis que fundamenta la pauta opera mediante una articulación vertical. En ella, cada nivel conceptual informa al siguiente: las categorías epistemológicas, como qualia e intencionalidad, se traducen en categorías fenomenológicas vinculadas con presencia y alteridad digital; estas se proyectan en categorías semióticas, como representamen, connotación y función semiótica, y finalmente se reorganizan en criterios de diseño evaluables.

A su vez, la articulación horizontal pone en diálogo las distintas tradiciones en cada nivel, permitiendo que sus categorías se complementen y tensionen productivamente. El resultado es un modelo integrador que entiende la percepción de consciencia como un fenómeno emergente entre disposición del usuario, experiencia de interacción, configuración semiótica de la interfaz y condiciones de diseño, sistematizado en la pauta final que se presenta en la Tabla 32.

Esta organización permite que la pauta no opere como una medición de consciencia real, sino como un instrumento de análisis de las condiciones que hacen posible su percepción. Por ello, las dimensiones no describen atributos internos del *chatbot*, sino condiciones e indicios observables en la interfaz, el discurso y las dinámicas de interacción.

Tabla 32

Pauta Final de evaluación metodológica para la percepción de consciencia en chatbots inteligentes

Dimensión	N.°	Criterio / categorías	Indicadores observables	Registro cualitativo sugerido
D1. **Agencia percibida**	1	Indistinguibilidad conductual y consciencia atribuida	El agente produce respuestas que pueden interpretarse como intencionales, pertinentes o deliberadas. Se distingue entre consciencia genuina y consciencia atribuida.	Baja / media / alta. Registrar evidencia textual o visual.
	2	Señales de intencionalidad en la interfaz	Avatar, nombre, tono verbal, *microcopys* o respuestas connotan propósito, personalidad o iniciativa.	Baja / media / alta. Describir signos que activan la atribución.
	3	Transparencia y límites del sistema	El agente declara qué puede y no puede hacer, advierte sobre errores y evita inducir al usuario a atribuciones falsas.	Insuficiente / parcial / adecuada. Registrar formulaciones de límite.
D2. **Presencia interactiva**	4	Transparencia y ruptura de la herramienta	La interacción fluye sin fricción; cuando se interrumpe, la ruptura permite observar el límite entre herramienta y agente.	Fluida / intermitente / fragmentada. Registrar momentos de ruptura.

Dimensión	N.°	Criterio / categorías	Indicadores observables	Registro cualitativo sugerido
	5	Affordance temporal y señales de actividad	Se observan indicador de escritura, latencia proporcional, señales de carga, envío o confirmación que sugieren actividad del agente.	Ausente / parcial / consistente. Registrar señales temporales.
	6	Acoplamiento y fluidez conversacional	La conversación mantiene turnos claros, continuidad temática, ajuste al usuario y ritmo coherente con la tarea.	Bajo / medio / alto. Describir continuidad y acoplamiento.
D3. **Identidad percibida**	7	Identidad singular del agente	Avatar, nombre, tipografía, color, tono y estilo permiten reconocer una entidad diferenciable.	Débil / moderada / fuerte. Registrar signos de identidad.
	8	Coherencia visual, verbal y funcional	La identidad visual, el tono verbal y la función del agente son consistentes durante la interacción.	Incoherente / parcialmente coherente / coherente.
	9	Connotaciones de presencia social	El diseño activa asociaciones de cercanía, empatía, tecnicidad, autoridad, juego o acompañamiento.	Describir connotaciones dominantes y posibles tensiones.
D4. **Comprensión aparente**	10	Coherencia conversacional y continuidad temática	El agente retoma información previa, mantiene el hilo temático y responde de forma pertinente al contexto inmediato.	Baja / media / alta. Registrar evidencia conversacional.

Dimensión	N.°	Criterio / categorías	Indicadores observables	Registro cualitativo sugerido
	11	Reparación conversacional y manejo del error	El agente solicita aclaraciones, corrige malentendidos, reconoce límites y ofrece alternativas.	Insuficiente / parcial / adecuada. Registrar secuencias de reparación.
	12	Naturalidad discursiva y adecuación pragmática	Las respuestas son claras, proporcionales y situadas; las fórmulas de cortesía no inducen confusión sobre experiencia subjetiva.	Mecánica / funcional / natural. Registrar ejemplos.
	13	Control del usuario y cierre de secuencia	El usuario puede editar, corregir, regenerar, salir o cerrar la interacción; las tareas tienen inicio, desarrollo y cierre reconocibles.	Bajo / medio / alto. Registrar controles disponibles.
Criterio transversal	14	Accesibilidad, inclusión y legibilidad	La interfaz es perceptible, operable, comprensible y robusta; el lenguaje es adecuado a la audiencia y al contexto de uso.	Insuficiente / parcial / adecuada. Registrar barreras o facilitadores.

La columna de registro cualitativo orienta la organización e interpretación de la evidencia de cada caso conforme a los criterios de la pauta. En consecuencia, la pauta debe entenderse como un instrumento orientado a registrar indicios, tensiones y efectos de diseño vinculados con la percepción de consciencia. Para ejemplificar su aplicación, la Tabla 33 presenta una lectura demostrativa del caso ChatGPT.

Tabla 33

Aplicación demostrativa de la pauta final de evaluación metodológica al caso ChatGPT

N.°	Dimensión	Criterio / categorías	Evidencia observada en ChatGPT	Valoración cualitativa
1	D1. Agencia percibida	Indistinguibilidad conductual y consciencia atribuida	ChatGPT declara que no posee consciencia, pero responde con continuidad, coherencia y ajuste al contexto.	Alta agencia percibida con baja atribución de experiencia subjetiva.
2	D1. Agencia percibida	Señales de intencionalidad en la interfaz	La intención no se construye por avatar, sino por pertinencia discursiva, organización argumentativa y tono colaborativo.	Media-alta. La agencia se sostiene en el discurso.
3	D1. Agencia percibida	Transparencia y límites del sistema	El agente explicita que no siente, no percibe y puede cometer errores.	Adecuada. La transparencia funciona como contrapeso ético.
4	D2. Presencia interactiva	Transparencia y ruptura de la herramienta	La conversación fluye hasta que el sistema declara límites o rectifica expresiones empáticas.	Fluida con rupturas productivas.
5	D2. Presencia interactiva	Affordance temporal y señales de actividad	El indicador de escritura y la latencia proporcional sugieren procesamiento activo.	Consistente. Refuerza presencia sin afirmar consciencia.
6	D2. Presencia interactiva	Acoplamiento y fluidez conversacional	El agente adapta el tono, retoma información y responde sin interrumpir.	Alto acoplamiento conversacional.
7	D3. Identidad percibida	Identidad singular del agente	No hay avatar antropomorfo; la identidad emerge del estilo conversacional y del nombre de la plataforma.	Moderada. Identidad discursiva más que visual.

N.°	Dimensión	Criterio / categorías	Evidencia observada en ChatGPT	Valoración cualitativa
8	D3. Identidad percibida	Coherencia visual, verbal y funcional	La interfaz austera es coherente con el tono técnico-colaborativo del agente.	Coherente. Baja humanización gráfica, alta estabilidad funcional.
9	D3. Identidad percibida	Connotaciones de presencia social	Las fórmulas de cortesía y el tono colaborativo activan cercanía, aunque el sistema corrige atribuciones emocionales.	Moderada. Presencia social regulada por transparencia.
10	D4. Comprensión aparente	Coherencia conversacional y continuidad temática	El agente mantiene el hilo de la conversación y organiza respuestas pertinentes.	Alta comprensión aparente.
11	D4. Comprensión aparente	Reparación conversacional y manejo del error	Ante cuestionamientos, rectifica y explica el uso de expresiones empáticas como fórmulas discursivas.	Adecuada. La reparación refuerza la comprensión aparente.
12	D4. Comprensión aparente	Naturalidad discursiva y adecuación pragmática	El discurso es claro y proporcionado; algunas fórmulas empáticas pueden producir ambigüedad si no son corregidas.	Natural con riesgo puntual de sobreatribución.
13	D4. Comprensión aparente	Control del usuario y cierre de secuencia	La interfaz permite editar, regenerar y cerrar. La conversación puede finalizar con despedida explícita.	Alto control del usuario.
14	Criterio transversal	Accesibilidad, inclusión y legibilidad	Tipografía legible, interfaz limpia y estructura conversacional comprensible. No se evaluaron usuarios con diversidad funcional.	Adecuada en observación general; requiere validación específica para accesibilidad.

Nota. Elaboración propia a partir de la PoC aplicada a ChatGPT.

La aplicación demostrativa ilustra el uso del instrumento, pero no reemplaza una evaluación con usuarios. Su función es mostrar cómo las dimensiones definitivas pueden organizar la evidencia registrada en los casos analizados sin volver a estructurarse por pilares.

En cuanto a las consideraciones metodológicas sobre la aplicación de la pauta, es necesario aclarar que no es un instrumento de medición cuantitativa ni una lista de cotejo de verificación binaria (sí/no). Es un dispositivo analítico cualitativo que opera mediante la descripción e interpretación de indicadores en cada dimensión, articulando los resultados del análisis documental con los datos emergentes de las entrevistas a expertos y del análisis de las pruebas de concepto (PoC). Su validez no descansa en la exhaustividad empírica, sino en la coherencia interna del marco teórico que la sustenta y en la capacidad de los indicadores para capturar el fenómeno de la percepción de consciencia con independencia del avance tecnológico de las plataformas analizadas.

En síntesis, el producto final de la investigación es una pauta de evaluación de la percepción de consciencia en chatbots inteligentes, estructurada en dimensiones, categorías, subcategorías e indicadores, con cuatro características: transferibilidad, escalabilidad, fundamentación disciplinar y operatividad. En este sentido, la tesis produce conocimiento de segundo orden: no solo describe un fenómeno, sino que construye las condiciones para que ese fenómeno pueda ser descrito de manera sistemática en contextos futuros.

CONCLUSIONES

Las conclusiones se organizan según cinco criterios: examen de las hipótesis, cumplimiento de los objetivos específicos, balance del modelo teórico, transferencia del conocimiento y futuras líneas de investigación.

La hipótesis general se corrobora con matices: encuentra apoyo conceptual y cualitativo en los resultados de esta investigación, dentro de los límites del diseño empleado y del alcance definido para esta tesis doctoral. La articulación del análisis documental, las entrevistas y la prueba de concepto permitió desarrollar una contribución original al diseño de interacción: una propuesta de pauta fundamentada para examinar las condiciones que intervienen en la percepción de consciencia en *chatbots*.

La perspectiva de los usuarios se incorporó de manera indirecta mediante el análisis de las investigaciones empíricas recogidas en el estado de la cuestión, cuyos resultados aportaron evidencia sobre las relaciones entre señales de interfaz, antropomorfismo y atribución de capacidades mentales. En conjunto, estas fuentes ofrecen sustento al planteamiento de que la percepción de consciencia no depende exclusivamente de las capacidades tecnológicas del sistema, sino también de las condiciones de diseño y de las señales y dinámicas de interacción. Esta integración de evidencia empírica previa fundamenta el análisis desarrollado, aunque no constituye una evaluación directa de la pauta propuesta con usuarios.

Las tres primeras hipótesis de trabajo también se corroboran, en el plano conceptual e interpretativo, dentro de los límites del diseño empleado. La primera encuentra sustento en la comprensión de la consciencia percibida como una atribución emergente en la interacción; la segunda, en la identificación de los elementos gráficos, interactivos y conversacionales como mediadores de la intencionalidad atribuida; y la tercera, en la articulación de categorías complementarias provenientes de la semiótica visual, el affordance y las heurísticas de interfaz. La cuarta hipótesis de trabajo recibe apoyo conceptual con matices: la aplicación de las categorías en los casos seleccionados ofrece una base preliminar para su uso transversal, pero no demuestra la estabilidad de los criterios entre generaciones tecnológicas y contextos. Este aspecto permanece

abierto al contraste empírico posterior. La prueba de concepto permitió ilustrar la aplicación de las categorías y fundamentar su organización en la pauta propuesta, sin establecer efectos causales sobre la percepción de usuarios. En particular, el análisis de ChatGPT permitió distinguir entre señales de agencia y afirmaciones sobre experiencia subjetiva, lo que sustenta la separación de ambas dimensiones en el instrumento.

En cuanto al cumplimiento de los objetivos específicos, todos fueron logrados dentro del alcance definido para la investigación. El primero, referido al establecimiento del posicionamiento epistemológico, se cumple mediante la formulación del constructo de consciencia percibida y la distinción entre consciencia genuina y consciencia atribuida. El segundo, orientado a articular la fenomenología de la experiencia con el diseño de interacción, se alcanza a través de las categorías de transparencia, alteridad digital y acoplamiento emergente. El tercero, centrado en identificar y sistematizar las categorías semióticas e interactivas, se desarrolla mediante la matriz documental, las entrevistas, la prueba de concepto y la reinterpretación de las heurísticas. Finalmente, el cuarto objetivo, correspondiente a la construcción de una propuesta de pauta de evaluación metodológica, se concreta en el instrumento final compuesto por cuatro dimensiones y un criterio transversal de accesibilidad.

El modelo teórico de los cuatro pilares resultó coherente y productivo como fundamento, pero no como estructura final del instrumento. Su sistematización en cuatro dimensiones evaluativas constituye el aporte metodológico central: las dimensiones recombinan los pilares en una estructura aplicable que conserva su fundamento conceptual sin reproducir la organización de las teorías de origen. Su principal fortaleza es organizar la evaluación mediante categorías que no se reducen a las prestaciones de una generación tecnológica específica. Esta orientación ofrece una base conceptual para su adaptación a distintos sistemas, cuya estabilidad y pertinencia deberán examinarse ante cambios tecnológicos y contextuales. Su principal límite es que describe condiciones de posibilidad y no cuantifica sus efectos, por lo que requiere instrumentos derivados para mediciones específicas.

Respecto de la transferencia del conocimiento a la comunidad científica y profesional, la pauta está diseñada para ser usada por investigadores y profesionales del diseño de interacción, UX y comunicación digital, con orientaciones que facilitan la comprensión de sus fundamentos filosóficos. Su estructura de lo general a lo específico (dimensiones, categorías, subcategorías e indicadores) ofrece posibilidades de adaptación: quien la aplique puede seleccionar las subcategorías pertinentes para su caso, incorporar indicadores propios de una plataforma y comparar agentes de naturaleza distinta. Su articulación con metodologías de diseño centrado en el usuario y con enfoques cualitativos de análisis fundamenta su potencial aplicación en contextos educativos, comerciales, sanitarios, institucionales o asistenciales. Además, la comunidad de referencia incluye a la academia del diseño en Iberoamérica, en consonancia con la orientación del Doctorado en Diseño de la Universidad de Palermo.

La originalidad de esta investigación reside en la articulación de fundamentos semióticos, fenomenológicos y de diseño de interacción en una pauta para examinar la percepción de consciencia en chatbots. Su potencial disruptivo se sitúa en el desplazamiento de una explicación centrada exclusivamente en las capacidades tecnológicas hacia una perspectiva que considera la interfaz y la interpretación del usuario como componentes constitutivos de dicha percepción. Este aporte cuestiona la equivalencia entre sofisticación tecnológica y atribución de capacidades mentales, sin anticipar efectos transformadores que todavía requieren evaluación.

La decisión de no formular conclusiones empíricas generalizables responde al alcance de la investigación y al resguardo ético frente a un fenómeno que involucra atribuciones de capacidades mentales, expectativas y relaciones con sistemas artificiales. La evaluación de la estabilidad de los criterios propuestos y de sus variaciones culturales y contextuales requiere una agenda de investigación a largo plazo, posterior a esta tesis, con estudios complementarios, participación de equipos interdisciplinarios y recursos que permitan ampliar y diversificar las muestras. Estas condiciones exceden los recursos y el período de ejecución de esta tesis, pero no impiden cumplir su objetivo de construir y

fundamentar la propuesta. Su evaluación directa con usuarios y el examen longitudinal de sus criterios constituyen etapas posteriores de una línea de investigación proyectada a un horizonte de al menos diez años, orientada a establecer su desempeño, sus límites de aplicación y su adaptación ante cambios tecnológicos y contextuales.

Las futuras líneas de investigación comprenden al menos cuatro proyecciones relevantes: la evaluación empírica de la pauta mediante su aplicación con usuarios y el desarrollo y validación de escalas derivadas; su aplicación comparada a sistemas multimodales, de voz y avatares encarnados; la incorporación de mediciones afectivas o fisiológicas sugeridas por la computación afectiva; y el estudio diacrónico del vínculo usuario-agente a medida que los modelos evolucionan. A ellas puede sumarse el análisis del contexto sociocultural como posible modulador de la percepción, aspecto emergente en las entrevistas que la pauta incorpora, pero que requiere un desarrollo específico en futuras investigaciones.

En síntesis, la evidencia producida en esta tesis sustenta la construcción y fundamentación de la pauta propuesta como contribución conceptual y metodológica al diseño de interacción. Su evaluación directa con usuarios y el estudio de su transferencia constituyen la continuidad de este aporte en la agenda de investigación futura.

BIBLIOGRAFÍA

Fuentes primarias

Araujo, T. (2018). Living up to the chatbot hype: The influence of anthropomorphic design cues and communicative agency framing on conversational agent and company perceptions. *Computers in Human Behavior*, *85*, 183–189. https://doi.org/10.1016/j.chb.2018.03.051

Arguelles Toache, E., & Amaro Rosales, M. (2023). Preocupaciones éticas en el uso de inteligencia artificial, transparencia y derecho de acceso a la información. El caso de los chatbots en el gobierno de México, en el contexto de la COVID-19. *Estudios en Derecho a la Información*, (15), 85–111. https://doi.org/10.22201/iij.25940082e.2023.15.17472

Austin, J. L. (1962). *How to do things with words* (J. O. Urmson, Ed.). Oxford University Press.

Bakhtin, M. M. (1986). *Speech genres and other late essays* (C. Emerson & M. Holquist, Eds.; V. W. McGee, Trad.). University of Texas Press. https://doi.org/10.7560/720466

Barthes, R. (1977). *Image, music, text* (S. Heath, Trad.). Hill and Wang.

Beck, U., Giddens, A., & Lash, S. (1994). *Reflexive modernization: Politics, tradition and aesthetics in the modern social order*. Stanford University Press.

Biblioteca del Congreso Nacional de Chile. (2015, 14 de abril). *Adrián Palacios, en el IV Congreso del Futuro* [Video]. YouTube. https://youtu.be/3QTcvPsUG7E

Bonsiepe, G. (1999). *Del objeto a la interfase: Mutaciones del diseño*. Infinito.

Borsci, S., Malizia, A., Schmettow, M., van der Velde, F., Tariverdiyeva, G., Balaji, D., & Chamberlain, A. (2022). The chatbot usability scale: The design and pilot of a usability scale for interaction with AI-based conversational agents. *Personal and Ubiquitous Computing*, *26*(1), 95–119. https://doi.org/10.1007/s00779-021-01582-9

Brentano, F. (1973). *Psychology from an empirical standpoint* (O. Kraus & L. L. McAlister, Eds.; A. C. Rancurello, D. B. Terrell, & L. L. McAlister, Trads.). Humanities Press. (Trabajo original publicado en 1874).

Brown, T. (2008). Design thinking. *Harvard Business Review*, *86*(6), 84–92.

Cámara de Diputadas y Diputados de Chile. (2024). *Regula los sistemas de inteligencia artificial* (Boletín N.º 16.821-19). https://www.camara.cl/legislacion/ProyectosDeLey/tramitacion.aspx?prmID=17429&prmBOLETIN=16821-19

Cao, Y., Liu, Y., Lyu, W., & Li, M. (2025). Human or AI customer service? The role of anthropomorphic avatars in shaping first impressions of chatbots. *Human Factors and Ergonomics in Manufacturing & Service Industries*, *35*(5), artículo e70024. https://doi.org/10.1002/hfm.70024

Chalmers, D. J. (1995). Facing up to the problem of consciousness. *Journal of Consciousness Studies*, *2*(3), 200–219.

Chen, J., Guo, F., Ren, Z., Li, M., & Ham, J. (2024). Effects of anthropomorphic design cues of chatbots on users' perception and visual behaviors. *International Journal of Human-Computer Interaction*, *40*(14), 3636–3654. https://doi.org/10.1080/10447318.2023.2193514

Cheng, X., Zhang, X., Cohen, J., & Mou, J. (2022). Human vs. AI: Understanding the impact of anthropomorphism on consumer response to chatbots from the perspective of trust and relationship norms. *Information Processing & Management, 59(3)*, *102940*. https://doi.org/10.1016/j.ipm.2022.102940

Congreso Futuro. (2015). *Introducción panel: La encrucijada de la vida ¿Las máquinas y la IA nos superarán?* https://www.congresofuturo.cl/charlas/la-encrucijada-de-la-vida

Cross, N. (2007). *Designerly ways of knowing*. Birkhäuser.

de Souza, C. S. (2005). *The semiotic engineering of human-computer interaction*. MIT Press.

de Souza, R. L. L., Maciel, C., & Nunes, E. P. dos S. (2021). Inspeção semiótica no sistema do Metahuman Creator: Avatares em foco. En *Anais da XXI Escola Regional de Informática de Mato Grosso* (pp. 77–83). Sociedade Brasileira de Computação. https://doi.org/10.5753/eri-mt.2021.18228

Dennis, A. R., Kim, A., Raimi, R., & Ayabakan, S. (2020). User reactions to COVID-19 screening chatbots from reputable providers. *Journal of the American Medical Informatics Association*, *27*(11), 1727–1731. https://doi.org/10.1093/jamia/ocaa167

Eco, U. (1976). *A theory of semiotics*. Indiana University Press.

Epley, N., Waytz, A., & Cacioppo, J. T. (2007). On seeing human: A three-factor theory of anthropomorphism. *Psychological Review*, *114*(4), 864–886. https://doi.org/10.1037/0033-295X.114.4.864

Fairclough, N. (1995). *Critical discourse analysis: The critical study of language*. Longman.

Fernández Villa, M. F., Gutiérrez Ospina, A. T., & Arias Cortés, D. (2019). Aprendizaje de dos conceptos semióticos para el análisis cinematográfico. *Sophia*, *15*(1), 109–118. https://doi.org/10.18634/sophiaj.15v.1i.907

Følstad, A., Araujo, T., Law, E. L.-C., Brandtzaeg, P. B., Papadopoulos, S., Reis, L., Baez, M., Laban, G., McAllister, P., Ischen, C., Wald, R., Catania, F., Meyer von Wolff, R., Hobert, S., & Luger, E. (2021). Future directions for chatbot research: An interdisciplinary research agenda. *Computing*, *103*(12), 2915–2942. https://doi.org/10.1007/s00607-021-01016-7

Fogg, B. J. (2003). *Persuasive technology: Using computers to change what we think and do*. Morgan Kaufmann Publishers.

Foucault, M. (1972). *The archaeology of knowledge and the discourse on language* (A. M. Sheridan Smith, Trad.). Pantheon Books.

Frascara, J. (2000). *Diseño gráfico para la gente: Comunicaciones de masa y cambio social*. Infinito.

Frascara, J. (2004). *Communication design: Principles, methods, and practice*. Allworth Press.

Fuentes Barassi, C. (2011). *Conciencia e inteligencia artificial: Consideraciones críticas sobre la plausibilidad de que una máquina programada posea conciencia fenoménica* [Tesis de maestría, Universidad de Chile]. Repositorio Académico de la Universidad de Chile. https://repositorio.uchile.cl/handle/2250/108753

Gambino, A., Fox, J., & Ratan, R. A. (2020). Building a stronger CASA: Extending the computers are social actors paradigm. *Human-Machine Communication*, *1*, 71–86. https://doi.org/10.30658/hmc.1.5

Gaver, W. W. (1991). Technology affordances. En *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '91)* (pp. 79–84). Association for Computing Machinery. https://doi.org/10.1145/108844.108856

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

Gnewuch, U., Morana, S., & Maedche, A. (2017). Towards designing cooperative and social conversational agents for customer service. En *Proceedings of the International Conference on Information Systems (ICIS 2017)*. Association for Information Systems. https://aisel.aisnet.org/icis2017/HCI/Presentations/1/

Gnewuch, U., Morana, S., Adam, M. T. P., & Maedche, A. (2018). Faster is not always better: Understanding the effect of dynamic response delays in human-chatbot interaction. En *Proceedings of the 26th European Conference on Information Systems (ECIS 2018)*. Association for Information Systems. https://aisel.aisnet.org/ecis2018_rp/113/

Gnewuch, U., Morana, S., Adam, M. T. P., & Maedche, A. (2022). Opposing effects of response time in human-chatbot interaction. *Business & Information Systems Engineering*, *64*(6), 773–791. https://doi.org/10.1007/s12599-022-00755-x

Go, E., & Sundar, S. S. (2019). Humanizing chatbots: The effects of visual, identity and conversational cues on humanness perceptions. *Computers in Human Behavior*, *97*, 304–316. https://doi.org/10.1016/j.chb.2019.01.020

González Hernández, N. Y., & Victoria-Uribe, R. (2023). Diseño de experiencia de usuario para la transferencia de conocimientos en entornos de realidad virtual. *I+Diseño. Revista Científica de Investigación y Desarrollo en Diseño*, *18*, 83–96. https://doi.org/10.24310/idiseo.18.2023.17453

Google. (2024, 18 de septiembre). *Conversation design*. Google for Developers. https://developers.google.com/assistant/conversation-design/welcome

Graphic Design Festival Scotland. (2017). *Identidad visual GDFS 2017*. https://graphicdesignfestivalscotland.com/2017/2017-identity/

Gray, H. M., Gray, K., & Wegner, D. M. (2007). Dimensions of mind perception. *Science, 315*(5812), 619. https://doi.org/10.1126/science.1134475

Grice, H. P. (1975). Logic and conversation. En P. Cole & J. L. Morgan (Eds.), *Syntax and semantics: Vol. 3. Speech acts* (pp. 41–58). Academic Press.

Groupe µ. (1992). *Traité du signe visuel: Pour une rhétorique de l'image*. Seuil.

Guingrich, R. E., & Graziano, M. S. A. (2025). A longitudinal randomized control study of companion chatbot use: Anthropomorphism and its mediating role on social impacts. *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society, 8*(2), 1153. https://doi.org/10.1609/aies.v8i2.36618

Heidegger, M. (1962). *Being and time* (J. Macquarrie & E. Robinson, Trads.). Blackwell. (Trabajo original publicado en 1927).

Husserl, E. (2001). *Logical investigations* (D. Moran, Ed.; J. N. Findlay, Trad.; Vol. 2). Routledge. (Trabajo original publicado en 1901).

Kendig, C. E. (2016). What is proof of concept research and how does it generate epistemic and ethical categories for future scientific practice? *Science and Engineering Ethics*, *22*(3), 735–753. https://doi.org/10.1007/s11948-015-9654-0

Kress, G., & van Leeuwen, T. (2006). *Reading images: The grammar of visual design* (2.ª ed.). Routledge.

Laime Eulogio, S., & Rodriguez Quintana, M. (2023). La semiótica de color en el cine de autor. *Apuntes de Ciencia & Sociedad*, *11*(2), 45–57. https://doi.org/10.18259/acs.2023018

Langarano Guerrero, M., Montaluisa Yugla, F., & Navas Moya, M. (2022). Implementación de un chatbot con NLP para recibir pedidos en una plataforma de delivery. *Revista Tecnológica - ESPOL*, *34*(3), 157–170. https://doi.org/10.37815/rte.v34n3.958

Latour, B. (2007). *Nunca fuimos modernos: Ensayo de antropología simétrica* (V. Goldstein, Trad.). Siglo XXI Editores. (Obra original publicada en 1991)

Lemoine, B. (2022, 11 de junio). *What is LaMDA and what does it want?* Medium. https://cajundiscordian.medium.com/what-is-lamda-and-what-does-it-want-688632134489

Lewis, C. I. (1929). *Mind and the world order: Outline of a theory of knowledge*. Scribner's Sons.

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Sage.

Luger, E., & Sellen, A. (2016). “Like having a really bad PA”: The gulf between user expectation and experience of conversational agents. En *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (pp. 5286–5297). ACM. https://doi.org/10.1145/2858036.2858288

Maeda, J. (2006). *The laws of simplicity*. MIT Press.

Mafra, M. G. S., Nunes, K., Rocha, S., Braz Junior, G., Silva, A., Viana, D., Silva, W., & Rivero, L. (2024). Proposing usability-UX technologies for the design and evaluation of text-based chatbots. *Journal on Interactive Systems*, *15*(1), 234–251. https://doi.org/10.5753/jis.2024.3856

Maldonado, T. (1999). *Lo real y lo virtual* (A. L. Bixio, Trad.; 2.ª ed.). Gedisa.

Manovich, L. (2001). *The language of new media*. MIT Press.

Manovich, L. (2021). Interfaces for a global village: Nam June Paik, Marshall McLuhan, and the future. En *Next 15 minutes* [Catálogo de exposición]. Daejeon Museum of Art. https://manovich.net/content/04-projects/116-interfaces-for-a-global-village-nam-jun-paik-marshall-mcluhan-and-the-future/manovich.interfaces-for-global-villlage.2021.pdf

Matarrese, M., Cravino, A., Isidoro, M., & García de la Cárcova, A. (2019). *Documento de los posgrados de la Facultad de Diseño y Comunicación de la Universidad de Palermo* (Escritos en la Facultad N.º 148). Universidad de Palermo.

Maturana, H., & Varela, F. (1984). *El árbol del conocimiento: Las bases biológicas del entendimiento humano*. Editorial Universitaria.

McLuhan, M. (2003). *Understanding media: The extensions of man* (W. T. Gordon, Ed.; ed. crítica). Gingko Press. (Obra original publicada en 1964)

Merleau-Ponty, M. (2002). *Phenomenology of perception* (C. Smith, Trad.). Routledge. (Trabajo original publicado en 1945).

Ministerio de Ciencia, Tecnología, Conocimiento e Innovación. (2021a, 3 de diciembre). *Decreto 20: Aprueba Política Nacional de Inteligencia Artificial*. Biblioteca del Congreso Nacional de Chile. https://www.bcn.cl/leychile/navegar?idNorma=1169399&idVersion=2021-12-03

Ministerio de Ciencia, Tecnología, Conocimiento e Innovación. (2021b, 25 de octubre). *Ley N.º 21.383: Modifica la Carta Fundamental, para establecer el desarrollo científico y tecnológico al servicio de las personas*. Diario Oficial de la República de Chile. https://www.bcn.cl/leychile/navegar?idNorma=1166983

Ministerio de Ciencia, Tecnología, Conocimiento e Innovación. (2021c). *Política Nacional de Inteligencia Artificial*. Gobierno de Chile. https://minciencia.gob.cl/uploads/filer_public/bc/38/bc389daf-4514-4306-867c-760ae7686e2c/documento_politica_ia_digital_.pdf

Ministerio de Salud de Chile. (2021, 27 de septiembre). *COVID-19: Gobierno anuncia fin del Estado de Excepción*. https://www.minsal.cl/covid-19-gobierno-anuncia-fin-del-estado-de-excepcion/

Nadarzynski, T., Miles, O., Cowie, A., & Ridge, D. (2019). Acceptability of artificial intelligence (AI)-led chatbot services in healthcare: A mixed-methods study. *Digital Health*, *5*, Artículo 2055207619871808. https://doi.org/10.1177/2055207619871808

Nagel, T. (1974). What is it like to be a bat? *The Philosophical Review*, *83*(4), 435–450. https://doi.org/10.2307/2183914

Nass, C., & Moon, Y. (2000). Machines and mindlessness: Social responses to computers. *Journal of Social Issues*, *56*(1), 81–103. https://doi.org/10.1111/0022-4537.00153

Nass, C., Steuer, J., & Tauber, E. R. (1994). Computers are social actors. En *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '94)* (pp. 72–78). Association for Computing Machinery. https://doi.org/10.1145/191666.191703

Nordheim, C. B., Følstad, A., & Bjørkli, C. A. (2019). An initial model of trust in chatbots for customer service: Findings from a questionnaire study. *Interacting with Computers*, *31*(3), 317–335. https://doi.org/10.1093/iwc/iwz022

Norman, D. A. (2010). *El diseño de los objetos del futuro*. Paidós Ibérica.

Norman, D. A. (2013). *The design of everyday things* (ed. rev. y ampliada). Basic Books. (Obra original publicada en 1988)

Peirce, C. S. (2012). *Obra filosófica reunida* (N. Houser & C. Kloesel, Eds.; D. McNabb, Trad.; Vols. 1–2). Fondo de Cultura Económica.

Piaget, J. (1998). *La equilibración de las estructuras cognitivas: Problema central del desarrollo* (5.ª ed.). Siglo XXI Editores. (Obra original publicada en 1975)

Piletsky, E. (2019). Consciousness and unconsciousness of artificial intelligence. *Future Human Image*, *11*, 66–71. https://doi.org/10.29202/fhi/11/7

Prada Núñez, R., Avendaño Castro, W. R., & Hernández Suarez, C. A. (2022). Globalización y cultura digital en entornos educativos. *Revista Boletín Redipe*, *11*(1), 262–272. https://doi.org/10.36260/rbr.v11i1.1641

Prates, R. O., de Souza, C. S., & Barbosa, S. D. J. (2000). Methods and tools: A method for evaluating the communicability of user interfaces. *Interactions*, *7*(1), 31–38. https://doi.org/10.1145/328595.328608

Putnam, H. (1967). Psychological predicates. En W. H. Capitan & D. D. Merrill (Eds.), *Art, mind, and religion* (pp. 37–48). University of Pittsburgh Press.

Quijada, S., & Vergara, R. (2024). *Sobre la medición de la economía digital en Chile* (Puntos de Referencia N.º 695). Centro de Estudios Públicos. https://static.cepchile.cl/uploads/cepchile/2024/05/29-195436_mlao_pdr695_quijada.pdf

Raineri Bernain, G. R. (2021). *Diseño autopoiético: Iluminación y caracterización del diseño desde la teoría de la autopoiesis* [Tesis doctoral, Universidad de Palermo]. Universidad de Palermo. https://www.palermo.edu/dyc/doctorado_diseno/documentacion/Tesis_Raineri.pdf

Reeves, B., & Nass, C. (1996). *The media equation: How people treat computers, television, and new media like real people and places*. Cambridge University Press.

Rico Sesé, J. (2020). El diseñador gráfico en la era de la inteligencia artificial. *EME Experimental Illustration, Art & Design*, *8*(8), 66–73. https://doi.org/10.4995/eme.2020.13210

Sánchez Martínez, J. A. (2011). Cibercultura: Semiótica del avatar y los mundos virtuales. *Veredas. Revista del Pensamiento Sociológico*, *12*(22), 33–44. https://veredasojs.xoc.uam.mx/index.php/veredas/article/download/250/249/249

Sanders, E. B.-N., & Stappers, P. J. (2008). Co-creation and the new landscapes of design. *CoDesign*, *4*(1), 5–18. https://doi.org/10.1080/15710880701875068

Schön, D. A. (1983). *The reflective practitioner: How professionals think in action*. Basic Books.

Schwartz, R., Vassilev, A., Greene, K., Perine, L., Burt, A., & Hall, P. (2022). *Towards a standard for identifying and managing bias in artificial intelligence* (NIST Special Publication 1270). National Institute of Standards and Technology. https://doi.org/10.6028/NIST.SP.1270

Scolari, C. A. (2008). *Hipermediaciones: Elementos para una teoría de la comunicación digital interactiva*. Gedisa.

Scolari, C. A. (2018). *Las leyes de la interfaz: Diseño, ecología, evolución, tecnología*. Gedisa.

Searle, J. R. (1969). *Speech acts: An essay in the philosophy of language*. Cambridge University Press.

Searle, J. R. (1980). Minds, brains, and programs. *Behavioral and Brain Sciences*, *3*(3), 417–424. https://doi.org/10.1017/S0140525X00005756

Senado de la República de Chile. (2021, 24 de noviembre). *Protección de los neuroderechos: Inédita legislación va a la Sala*. https://www.senado.cl/comunicaciones/noticias/proteccion-de-los-neuroderechos-inedita-legislacion-va-la-sala

Shardlow, M., & Przybyła, P. (2022). *Deanthropomorphising NLP: Can a language model be conscious?* [Preprint]. arXiv. https://doi.org/10.48550/arXiv.2211.11483

Simon, H. A. (1996). *The sciences of the artificial* (3.ª ed.). MIT Press.

Simondon, G. (2018). *El modo de existencia de los objetos técnicos* (M. Martínez & P. Rodríguez, Trads.; 3.ª ed.). Prometeo Libros. (Obra original publicada en 1958)

Song, S. W., & Shin, M. (2024). Uncanny valley effects on chatbot trust, purchase intention, and adoption intention in the context of e-commerce: The moderating role of avatar familiarity. *International Journal of Human-Computer Interaction*, *40*(2), 441–456. https://doi.org/10.1080/10447318.2022.2121038

Stephenson, N. (1992). *Snow crash*. Bantam Books.

Suchman, L. A. (1985). *Plans and situated actions: The problem of human-machine communication* (ISL-6; Corporate Accession P85-00005). Xerox Corporation, Palo Alto Research Center.

Turing, A. M. (1950). Computing machinery and intelligence. *Mind*, *59*(236), 433–460. https://doi.org/10.1093/mind/LIX.236.433

Turkle, S. (2011). *Alone together: Why we expect more from technology and less from each other*. Basic Books.

Umair, M., Guliyeva, S., & Shakaraliyeva, Z. (2026). Between machine and human: How chatbot identity framing shapes consumer trust, comfort, and brand loyalty. *Strategic Business Research*. Publicación anticipada en línea. https://doi.org/10.1016/j.sbr.2026.100140

Universidad de Palermo. (s. f.). *Actualidad y devenir de los lenguajes visuales. Línea de Investigación N.º 10*. Facultad de Diseño y Comunicación. https://www.palermo.edu/dyc/coloquio_investigacion_y_desarrollo/comision umero10.html

Universidad de Santiago de Chile. (2019). *Resumen ejecutivo Informe de Evaluación Institucional*. https://usach.cl/sites/default/files/files/ResumenEjecutivoInformeV4.pdf

Universidad de Santiago de Chile. (2021). *Plan de fortalecimiento a 10 años Universidad de Santiago de Chile: Versión ajustada 2021*. https://usach.cl/sites/default/files/files/FIRMADO_USANTIAGO_PFE10A_OS_2021.pdf

van Dijk, T. A. (1980). *Macrostructures: An interdisciplinary study of global structures in discourse, interaction, and cognition*. Lawrence Erlbaum Associates.

van Dijk, T. A. (2008). *Discourse and context: A sociocognitive approach*. Cambridge University Press.

Varela, F. J., Thompson, E., & Rosch, E. (1991). *The embodied mind: Cognitive science and human experience*. MIT Press.

Vivanco, G., & Gorostiaga, J. M. (2017). Cultura digital y diversidad: Perspectivas de discursos de políticas TIC-educación. *Cadernos de Pesquisa*, *47*(165), 1016–1043. https://doi.org/10.1590/198053144261

von Glasersfeld, E. (1984). An introduction to radical constructivism. En P. Watzlawick (Ed.), *The invented reality: How do we know what we believe we know?* (pp. 17–40). W. W. Norton.

Waytz, A., Cacioppo, J. T., & Epley, N. (2010). Who sees human? The stability and importance of individual differences in anthropomorphism. *Perspectives on Psychological Science*, *5*(3), 219–232. https://doi.org/10.1177/1745691610369336

Weizenbaum, J. (1966). ELIZA-A computer program for the study of natural language communication between man and machine. *Communications of the ACM*, *9*(1), 36–45. https://doi.org/10.1145/365153.365168

Weizenbaum, J. (1976). *Computer power and human reason: From judgment to calculation*. W. H. Freeman and Company.

Westerman, D., Cross, A. C., & Lindmark, P. G. (2019). I believe in a thing called bot: Perceptions of the humanness of “chatbots”. *Communication Studies*, *70*(3), 295–312. https://doi.org/10.1080/10510974.2018.1557233

Winograd, T., & Flores, F. (1986). *Understanding computers and cognition: A new foundation for design*. Ablex Publishing Corporation.

World Wide Web Consortium. (2023, 5 de octubre). *Web content accessibility guidelines (WCAG) 2.2*. https://www.w3.org/TR/2023/REC-WCAG22-20231005/

Fuentes secundarias

Adamopoulou, E., & Moussiades, L. (2020). An overview of chatbot technology. En I. Maglogiannis, L. Iliadis, & E. Pimenidis (Eds.), *Artificial intelligence applications and innovations. AIAI 2020* (Vol. 584, pp. 373–383). Springer. https://doi.org/10.1007/978-3-030-49186-4_31

Alemany, L. (2021, 21 de junio). Se necesitan lingüistas y psicólogos para enseñar español a la Inteligencia Artificial. *El Mundo*. https://www.elmundo.es/cultura/2021/06/21/60d073cafdddff8a248b45a8.html

BBC News Mundo. (2022, 13 de junio). *El ingeniero de Google que asegura que un programa de inteligencia artificial ha cobrado conciencia propia y siente*. https://www.bbc.com/mundo/noticias-61787944

Berrones Yaulema, L. P., & Buenaño Barreno, P. N. (2023). ChatGPT en el ámbito educativo. *Esprint Investigación*, *2*(2), 45–54. https://doi.org/10.61347/ei.v2i2.57

Cihodaru-Ștefanache, Ș., & Podina, I. R. (2025). The uncanny valley effect in embodied conversational agents: A critical systematic review of attractiveness, anthropomorphism, and uncanniness. *Frontiers in Psychology, 16*, 1625984. https://doi.org/10.3389/fpsyg.2025.1625984

Cobarrubias Soto, N., Rodríguez López, H., & Zaldívar Colado, X. (2022). Diseño centrado en el usuario: Interacción hombre-computadora. *Revista Digital de Tecnologías Informáticas y Sistemas*, *6*(1), 78–84. https://doi.org/10.61530/redtis.2022.6.6.124.7

Cordeiro Costa, W., & Perassi Luiz de Sousa, R. (2022). Atualização do conhecimento em arte na cibersociedade. *Anais do Congresso Internacional de Conhecimento e Inovação (CIKI)*, *1*(1). https://doi.org/10.48090/ciki.v1i1.1145

Díaz, F. J., Harari, I., & Amadeo, A. P. (2013). *Guía de recomendaciones para diseño de software centrado en el usuario*. Editorial de la Universidad Nacional de La Plata. https://doi.org/10.35537/10915/32172

Elias, M. A., Faversani, L. A., Moreira, J. A. V., Masiero, A. V., & Cunha, N. V. da. (2023). Inteligencia artificial en salud y sus implicaciones bioéticas: Una revisión sistemática. *Revista Bioética*, *31*, Artículo e3542ES. https://doi.org/10.1590/1983-803420233542ES

Følstad, A., & Brandtzæg, P. B. (2017). Chatbots and the new world of HCI. *Interactions*, *24*(4), 38–42. https://doi.org/10.1145/3085558

Jenneboer, L., Herrando, C., & Constantinides, E. (2022). The impact of chatbots on customer loyalty: A systematic literature review. *Journal of Theoretical and Applied Electronic Commerce Research*, *17*(1), 212–229. https://doi.org/10.3390/jtaer17010011

Klinkenberg, J.-M. (2006). *Manual de semiótica general* (G. Baquero Heredia, Trad.). Universidad de Bogotá Jorge Tadeo Lozano.

Logan, R. K. (2010). *Understanding new media: Extending Marshall McLuhan*. Peter Lang.

Malpica Velasco, J. A. (s. f.). *Inteligencia artificial y conciencia*. Universidad de Alcalá. https://frasca.web.uah.es/inteligencia-artificial.pdf

Miranda Márquez, A. (2020). Inteligencia artificial demasiado humana (aún): Arte y tecnología. *Revista Digital Universitaria*, *21*(1). https://doi.org/10.22201/codeic.16076079e.2020.v21n1.a7

Moya López, I. A. (2009, noviembre). *Inteligencia artificial*. Universidad Mayor de San Andrés. https://inf166.wordpress.com/wp-content/uploads/2014/04/inteligencia-artificial.pdf

Nielsen, J. (1994). *Usability engineering*. Morgan Kaufmann.

Nielsen, J. (2024, 30 de enero). *10 usability heuristics for user interface design*. Nielsen Norman Group. https://www.nngroup.com/articles/ten-usability-heuristics/

Nielsen, J., & Budiu, R. (2013). *Usabilidad en dispositivos móviles*. Anaya Multimedia.

Nöth, W. (Ed.). (1997). *Semiotics of the media: State of the art, projects, and perspectives*. De Gruyter Mouton. https://doi.org/10.1515/9783110803617

Pandey, S. C. (2018). Can artificially intelligent agents really be conscious? *Sādhanā*, *43*(7), Artículo 110. https://doi.org/10.1007/s12046-018-0887-x

Piccolo, L. S. G., Mensio, M., & Alani, H. (2019). Chasing the chatbots: Directions for interaction and design research. En S. S. Bodrunova, O. Koltsova, A. Følstad, H. Halpin, P. Kolozaridi, L. Yuldashev, A. Smoliarova, & H. Niedermayer (Eds.), *Internet science* (pp. 157–169). Springer. https://doi.org/10.1007/978-3-030-17705-8_14

Rapp, A., Curti, L., & Boldi, A. (2021). The human side of human-chatbot interaction: A systematic literature review of ten years of research on text-based chatbots. *International Journal of Human-Computer Studies, 151*, 102630. https://doi.org/10.1016/j.ijhcs.2021.102630

Revilla Muñoz, O., & Carreras Montoto, O. (2025). *Web accessibility: WCAG 2.2 made easy*. Itákora Press.

Shneiderman, B. (1987). *Designing the user interface: Strategies for effective human-computer interaction*. Addison-Wesley.

Shneiderman, B., & Plaisant, C. (2005). *Designing the user interface: Strategies for effective human-computer interaction* (4.ª ed.). Pearson/Addison-Wesley.

Silva, M. da, Bonin, J. C., & Garrote, R. (2023). Elementos da cultura digital para o ensino de filosofia no Ensino Médio: O que dizem as pesquisas? *Educação em Análise*, *8*(1), 62–78. https://doi.org/10.5433/1984-7939.2023v8n1p62

Sonesson, G. (1994). Pictorial semiotics, Gestalt theory, and the ecology of perception. *Semiotica*, *99*(3–4), 319–400. https://doi.org/10.1515/semi-1994-993-405

Steven, R. (2017, 22 de junio). *Create a poster for Graphic Design Festival Scotland*. Creative Review. https://www.creativereview.co.uk/work/graphic-design-festival-scotland-identity-branding/

Tiku, N. (2022, 11 de junio). The Google engineer who thinks the company's AI has come to life. *The Washington Post*. https://www.washingtonpost.com/technology/2022/06/11/google-ai-lamda-blake-lemoine/

Vargas Márquez, B. L., Inga Hanampa, L. A., & Maldonado Portilla, M. G. (2021). Design thinking aplicado al diseño de experiencia de usuario. *Innovación y Software*, *2*(1), 6–19. https://doi.org/10.48168/innosoft.s5.a35

Yin, R. K. (2014). *Case study research: Design and methods* (5.ª ed.). Sage Publications.

Ynoub, R. (2015). *Cuestión de método: Aportes para una metodología crítica* (Tomo I). Cengage Learning.