

# INTERACTUANDO CON LA IA: EL PROBLEMA DEL VALLE INQUIETANTE Y LOS ESTILOS DE LENGUAJE QUE UTILIZAN LOS CHATBOTS

Desde nuestros inicios como sociedad moderna, los humanos hemos buscado crear tecnología que automatice o reemplace nuestras tareas. Ha sido inherente a ese proceso el que diseñemos robots que se parezcan física y cognitivamente a nosotros. Este deseo se ve representado claramente en películas de Hollywood. En muchas de ellas se muestran robots que pueden hablar, caminar y actuar como si fueran humanos. Los avances tecnológicos actuales están convirtiendo nuestra imaginación en realidad, con robots que ahora guían a turistas en aeropuertos (Joose & Evers, 2017), ayudan a los médicos a realizar cirugías (Berlyand et al., 2018) y asisten a meseros en restaurantes (Yang & Chew, 2021).

Los robots humanoides ofrecen una ventaja importante sobre los robots no humanoides, ya que "proporcionan una sensación de familiaridad y ayudan a los humanos a empatizar

con las máquinas" (Wagner, 2017). Sin embargo, existe una desventaja en los robots humanoides: cuanto más se parecen a los humanos, más tienden a resultar inquietantes y a veces repulsivos, un fenómeno conocido como "valle inquietante". Masahiro Mori, profesor de robótica en el Instituto Tecnológico de Tokio, escribió un ensayo (Mori, 1970) sobre cómo imaginaba las reacciones de las personas ante robots que parecieran y actuaran casi como un humano. En particular, planteó la hipótesis de que la respuesta de una persona hacia un robot humanizado cambiaría bruscamente de empatía a repulsión a medida que se acercara, pero sin lograr, a una apariencia más realista (Figura 1). Hay varios factores que contribuyen a estos sentimientos de incomodidad, incluyendo anomalías en características de baja calidad en robots humanoides, como un color de piel antinatural y ojos anormales, su parecido con personas muertas y la violación de las expectativas sobre cómo deben estar dispuestas las características faciales.

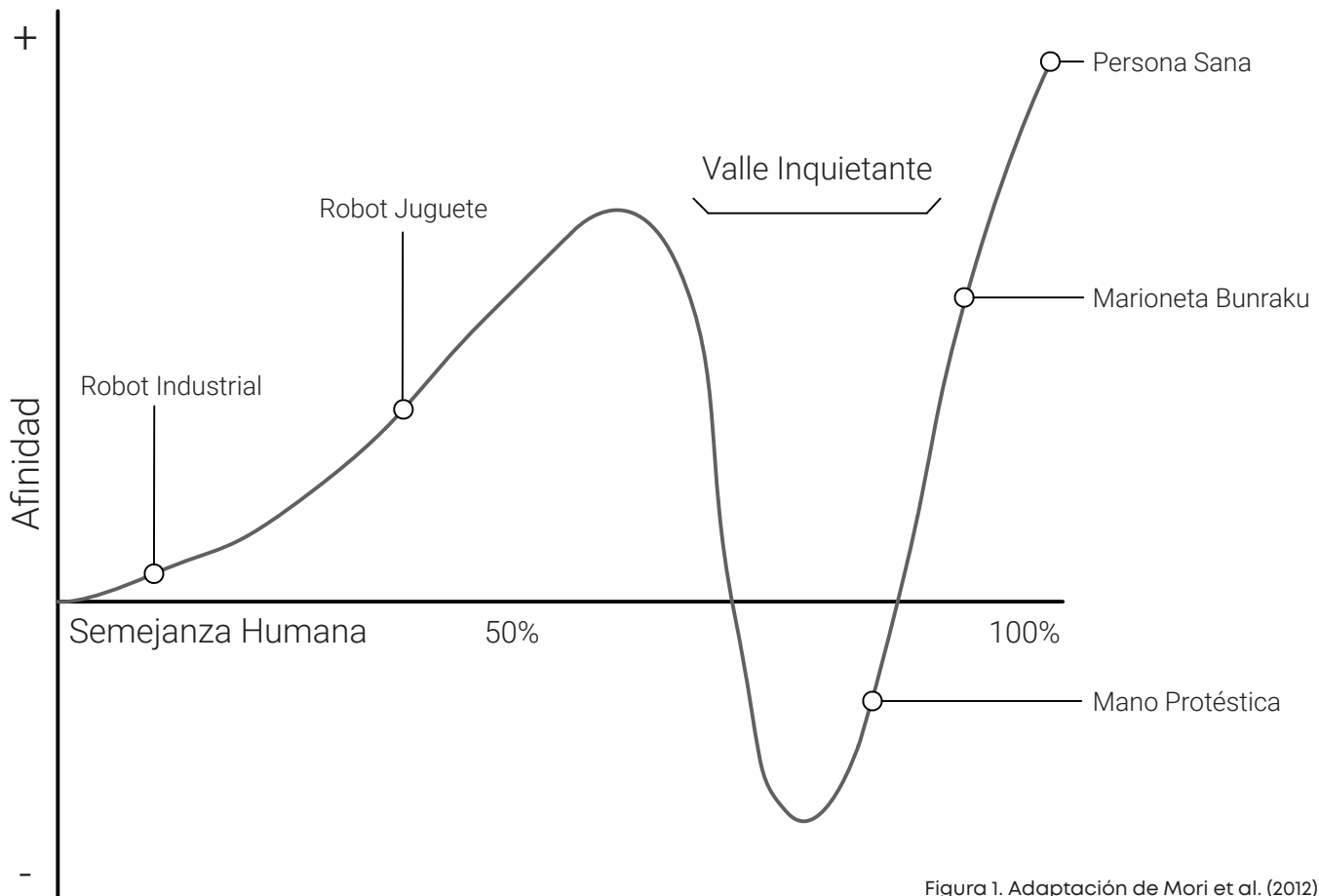


Figura 1. Adaptación de Mori et al. (2012)



Una explicación para el valle inquietante es que los robots humanoides son perturbadores porque parecen poseer una mente similar a la humana. Aunque los robots no tengan realmente una mente humana, se ha sugerido que a menudo se perciben mentes en ellos (Wang & Krumhuber, 2018). La percepción de si alguien o algo tiene una mente suele depender del observador. Gray and Wegner (2012) proponen que las personas perciben las mentes a lo largo de dos dimensiones: agencia y experiencia. La agencia se refiere a la capacidad de pensar, planificar y actuar, mientras que la experiencia es la capacidad de sentir emociones y sensaciones corporales, como hambre, placer y dolor.

Estudios revelan que las percepciones de experiencia se ven tanto más esencialmente presentes en los humanos como esencialmente ausentes en las máquinas (Tzelios et al., 2022). Por esta razón, un robot que es percibido como capaz de mostrar aspectos de la dimensión de experiencias, es decir, la capacidad de sentir emociones, podría generar percepciones dentro del valle inquietante. Esta hipótesis ha sido confirmada por una variedad de estudios (Gray & Wegner, 2012; Shank et al., 2019). Los robots resultan inquietantes cuando se perciben como capaces de tener emociones humanas, ya sea por su apariencia, sus acciones o simplemente por cómo se describen.

La explicación de la percepción de la mente en el valle inquietante sugiere que las características humanas de los robots humanoides, que pueden ser inquietantes por sí mismas, también llevan a las personas a ver mentes humanas, que involucran esencialmente la capacidad de sentir y tener experiencias conscientes. Aunque los humanos pueden percibir espontáneamente algunos niveles de agencia en los robots, percibir la experiencia en robots entra en conflicto con el conocimiento explícito de las personas de que los robots, como simples máquinas, carecen de estas capacidades mentales humanas, y esta discrepancia crea una aversión afectiva.

El concepto de percepción mental en relación con el valle inquietante sostiene que las características humanas de los

robots humanoides no solo causan malestar, sino que también llevan a las personas a atribuirles estados mentales similares a los humanos, principalmente la capacidad de sentir y ser conscientes (Schein & Gray, 2015). Aunque los humanos naturalmente asignan algún grado de autonomía a los robots, reconocer cualidades experienciales en ellos entra en conflicto con el claro entendimiento de que, como máquinas simples, no poseen estos rasgos mentales humanos. Esta incongruencia resulta en malestar emocional (Gray & Wegner, 2012; Shank et al., 2019).

La reciente popularidad de herramientas como ChatGPT y Bard ilustra la creciente presencia de la IA en la creación de contenido en línea. Este fenómeno refleja nuestra continua fascinación por crear entidades que nos imiten, similar a los avances en robots humanoides que asisten en diversas áreas, desde la medicina hasta el hospitality. Sin embargo, al igual que el Valle Inquietante sugiere una reacción emocional adversa hacia robots que casi replican la apariencia humana, la IA en el marketing digital enfrenta un desafío parecido: cómo ser efectiva sin causar rechazo.

Los usuarios que revelan el uso de IA en su contenido buscan demostrar vanguardia tecnológica, pero también deben navegar el equilibrio delicado de humanizar la IA sin cruzar hacia el Valle Inquietante. Esto resalta la necesidad de comprender los efectos de la IA en la percepción de los consumidores y los factores que modifican estos efectos.

Las compañías que deciden hacer explícito su uso de bots al crear su contenido y respuestas piensan que esta práctica será bien recibida por su audiencia. Crean que demostrará que están a la vanguardia de la tecnología y que compartir la autoría protegerá su proceso creativo y sus resultados (Suh et al., 2021). De esta manera, será importante comprender más profundamente los efectos de esta nueva tendencia en los comportamientos y percepciones de los consumidores, los caminos que explican estos efectos y cómo varios factores podrían alterarlos.

Una consideración relevante para entender este efecto es que, aunque el canal de comunicación puede ser dinámico, el mensajero y el mensaje deben ser coherentes, tanto en sustancia como en forma (por ejemplo, Guy, 2013). En este sentido, la Teoría de la Expectativa del Lenguaje propone que los receptores generan estereotipos sobre los mensajeros y esperan que el mensaje esté alineado con ese

estereotipo (Dillard & Pfau, 2002). Por lo tanto, si hay inconsistencias entre el mensajero y el mensaje, el mensaje no será necesariamente bien recibido. En el contexto de nuestro estudio, esto significa que, si las personas saben que un bot genera el contenido, esperarán que el mensaje muestre características típicas de una máquina. Esta expectativa se alinea con el concepto del Valle Inquietante, donde la semejanza excesiva con los humanos en los robots puede ser perturbadora. Por lo tanto, un lenguaje claramente 'no humano' podría ser más efectivo y menos inquietante en la comunicación de marca mediante IA.

Los resultados preliminares de nuestro estudio indican que los usuarios pasan por un valle inquietante cuando un bot contesta de forma demasiado humana, y eso hace que la percepción sobre el mensaje se vea afectado. Esto es interesante a nivel de marketeros que se encuentran implementando sistemas de IA en las conversaciones que mantienen con sus clientes online. Muchos deciden ocultar que están ocupando un bot y deciden también incluir una foto de un humano con nombre propio. Esto podría causar efectos negativos al momento que el cliente descubra que está interactuando con un bot, más aún si este bot ha utilizado un lenguaje que sea considerado como extremadamente humano. ▲

## REFERENCIAS

Berlyand, Y., Raja, A. S., Dörner, S. C., Prabhakar, A. M., Sonis, J. D., Gottumukkala, R. V., Succu, M. D., & Yun, B. J. (2018). How artificial intelligence could transform emergency department operations. *Am J Emerg Med*, 36(8), 1515-1517. <https://doi.org/10.1016/j.ajem.2018.01.017>.

Dillard, J., & Pfau, M. (2002). The Persuasion Handbook: Developments in Theory and Practice. <https://doi.org/10.4135/9781412976046>

Gray, K., & Wegner, D. M. (2012). Feeling robots and human zombies: mind perception and the uncanny valley. *Cognition*, 125(1), 125-130. <https://doi.org/10.1016/j.cognition.2012.06.007>

Guy, G. R. (2013). The cognitive coherence of sociolects: How do speakers handle multiple sociolinguistic variables? *Journal of Pragmatics*, 52, 63-71. <https://doi.org/10.1016/j.pragma.2012.12.019>

Joose, M., & Evers, V. (2017). A Guide Robot at the Airport Proceedings of the Companion of the 2017 ACM/IEEE International Conference on Human-Robot Interaction.

Mori, M. (1970). The uncanny valley. *Energy*, 7, 33-35, Article 4.

Mori, M., MacDorman, K., & Kageki, N. (2012). The Uncanny Valley [From the Field]. *IEEE Robotics & Automation Magazine*, 19(2), 98-100. <https://doi.org/10.1109/mra.2012.2192811>

Schein, C., & Gray, K. (2015). The eyes are the window to the uncanny valley. *Interaction Studies. Social Behaviour and Communication in Biological and Artificial Systems*, 16(2), 173-179. <https://doi.org/10.1075/is.16.2.02sch>

Shank, D. B., Graves, C., Gott, A., Gamez, P., & Rodriguez, S. (2019). Feeling our way to machine minds: People's emotions when perceiving mind in artificial intelligence. *Computers in Human Behavior*, 98, 256-266. <https://doi.org/10.1016/j.chb.2019.04.001>

Suh, M., Youngblom, E., Terry, M., & Cai, C. J. (2021). AI as Social Glue: Uncovering the Roles of Deep Generative AI during Social Music Composition Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems.

Tzelios, K., Williams, L. A., Omerod, J., & Bliss-Moreau, E. (2022). Evidence of the unidimensional structure of mind perception. *Sci Rep*, 12(1), 18978. <https://doi.org/10.1038/s41598-022-23047-6>

Wagner, A. (2017). Why do we keep making robots that look like people? This video has some answers. *Science*. <https://doi.org/10.1126/science.aar8229>

Wang, X., & Krumhuber, E. G. (2018). Mind Perception of Robots Varies With Their Economic Versus Social Function. *Front Psychol*, 9, 1230. <https://doi.org/10.3389/fpsyg.2018.01230>

Yang, J., & Chew, E. (2021). The Design Model for Robotic Waitress. *International Journal of Social Robotics*, 13(7), 1541-1551. <https://doi.org/10.1007/s12369-021-00745-1>.



**Felix Lizama**

PhD in Information Technology  
Profesor Asistente Departamento de  
Administración Facultad de  
Economía y Negocios  
**Universidad de Chile**