

<https://doi.org/10.1038/s44271-025-00376-6>

# Metaphors of AI indicate that people increasingly perceive AI as warm and human-like

Check for updates

Myra Cheng<sup>1,4</sup> , Angela Y. Lee<sup>1,2,4</sup> , Kristina Rapuano<sup>3</sup>, Kate Niederhoffer<sup>3</sup>, Alex Liebscher<sup>3</sup> & Jeffrey Hancock<sup>1</sup>

As AI-based technologies such as ChatGPT are increasingly used across various sectors, understanding how people conceptualize artificial intelligence (AI) is crucial for anticipating public response and developing AI technologies responsibly<sup>1</sup>. We hypothesize that public perceptions of AI are rapidly evolving, and that these perceptions inform not only how people use AI, but also the extent to which they trust it and the role they believe it should play in their lives - if at all. However, beliefs about complex sociotechnical systems like AI are nuanced and hard to articulate<sup>2-4</sup>, especially using traditional self-reporting methods where people may struggle to clearly articulate their implicit attitudes about emerging technologies<sup>5</sup>. To overcome these limitations, we collected over 12,000 open-ended metaphor responses over 12 months from a nationally representative U.S. sample and developed a systematic framework to quantitatively analyze them. Here we show that US Americans perceive AI as warm and competent, with attributions of human-likeness and warmth increasing significantly in the year after ChatGPT was introduced, and that these perceptions strongly predict trust and willingness to adopt AI technologies. We also identify important demographic variations, with women, older individuals, and people of color more likely to attribute human-like qualities to AI, helping explain disparities in trust and adoption rates. This scalable metaphor analysis approach enables tracking multifaceted public attitudes to inform AI governance, revealing how perceptions influence technology adoption across different populations.

Advances in large language models (LLMs) have catalyzed widespread public interest in artificial intelligence (AI), with over 90% of US Americans having heard of AI and tools like ChatGPT receiving over a billion queries a day<sup>6</sup>. The integration of AI into many facets of life raises important questions about what this change means for society<sup>7,8</sup>. While some forecast AI to become a transformative tool for productivity<sup>9</sup> or a trusted assistant for personal tasks<sup>10</sup>, others express concerns. For instance, many fear that individual reliance on AI can impair critical thinking processes<sup>11</sup> and that organizational adoption of AI may replace jobs and threaten livelihoods<sup>12,13</sup>.

These perceptions are important determinants of crucial decisions, such as whether individuals will choose to use it, whether industries will adopt it, and whether the public will support or reject regulatory measures regarding AI<sup>1</sup>. For example, an individual's decision to use AI to answer their questions, process their emotions, or assist them with their work is highly dependent on their perception of the system's capabilities<sup>14</sup>.

Decisions to regulate AI are also fundamentally shaped by perceptions about these systems, such as whether they provide opportunities for growth or threats to human progress<sup>15</sup>. Understanding public perceptions of AI is therefore integral to tracking how society is responding to the widespread adoption of AI and anticipating whether the extent to which they trust it going forward.

Perceptions about new technologies, however, can be difficult to capture using traditional survey measures<sup>2,16,17</sup>. People often struggle to verbalize their nuanced perceptions of complex sociotechnical systems like AI, particularly when they have less experience with them<sup>18,19</sup>. We therefore complement existing, large-scale survey investigations into public perceptions of AI (e.g., refs. 20,21) by analyzing the metaphors people use to describe AI. People have long used metaphors to communicate complex ideas, like the turbulence of emotions (“rollercoaster of feelings”) or the value of digital resources (“data is the new oil”) <sup>22</sup>. The metaphors people use

<sup>1</sup>Stanford University, Stanford, CA, USA. <sup>2</sup>School of Journalism & Mass Communication, University of Wisconsin-Madison, Madison, USA. <sup>3</sup>BetterUp Labs, Austin, TX, USA. <sup>4</sup>These authors contributed equally: Myra Cheng, Angela Y. Lee. e-mail: [myra@cs.stanford.edu](mailto:myra@cs.stanford.edu); [angela8@stanford.edu](mailto:angela8@stanford.edu)

to understand AI (“as a personal assistant”, “as a tutor”) can reveal the rich, implicit, and nuanced ways they are responding to the recent widespread adoption of this technology.

We investigate the dynamics of US Americans’ perceptions of AI by collecting over 12,000 metaphorical descriptions of AI from a nationally representative U.S. sample between May 2023 and August 2024, a period that captured a major increase in public attention with generative AI given the introduction of ChatGPT<sup>23</sup>. We use computational methods to analyze this large-scale dataset and address three core research questions:

First, how do members of the public conceptualize AI, and how do these conceptualizations vary across demographics? Examining the metaphors people use to describe novel developments, like AI, through analogies to familiar concepts (“AI is a friendly teacher” vs. “AI is a powerful brain”) can provide insight into how different people are thinking and responding to societal changes<sup>24,25</sup>. Building on prior work indicating demographic differences in perceptions of AI (i.e., where women were more likely to express concern or distrust than men<sup>20,26,27</sup>), we examine how age, gender, race, education, and work experience shape the dominant metaphors people of different backgrounds use to understand AI.

Second, are public perceptions of AI evolving over time, and if so, in what ways? The temporal nature and large scale of our dataset enable us to examine shifts in how people are thinking about AI by assessing changes in their metaphors. While “AI” was once a relatively niche technical term<sup>28</sup>, it is increasingly associated with user-facing large-language models (LLMs) and chatbots with conversational interfaces<sup>21</sup>. Just as Steve Jobs once marketed the laptop as “a bicycle for the mind”, AI is increasingly marketed to the public as “assistants,” “agents,” “copilots,” or even “companions.” We examine if people draw on similar themes when asked to describe AI in their own words by using NLP methods to measure two types of implicit perceptions from our metaphors dataset: (1) anthropomorphism—the attribution of human-like qualities—which we measure using AnthroScore, a method for surfacing implicit framings via language model (LM)-based probabilities<sup>29</sup>, and (2) perceptions of AI as a social entity—specifically warmth and competence—using semantic axes constructed with LM-based embeddings<sup>30</sup>. Our findings allow us to examine whether people are viewing AI as more similar to humans, and if so, the extent to which they view them as friendly and capable.

Finally, can examinations of metaphors of AI help us understand who chooses to trust and adopt AI? Warmth and competence are integral dimensions of how people form impressions of social actors<sup>31,32</sup>. The extent to which people believe that others have warm, positive intentions and are capable of carrying out given tasks is a strong predictor of people’s attitudes, such as their trust in new technologies and willingness to adopt them in the future<sup>14,33–35</sup>. A broad literature has examined factors influencing trust in and adoption of AI, such as demographic characteristics, political orientation, and media exposure<sup>36</sup>. Building on prior findings that perceptions of warmth and competence predict trust and willingness to adopt AI<sup>37–40</sup>, we hypothesize that metaphors and implicit perceptions offer additional explanatory power.

Metaphors are powerful cognitive and linguistic tools that help people understand information by distilling complex concepts into more accessible ones through analogies<sup>24,41</sup>. Decades of psychological research demonstrate that the metaphors people use to understand abstract concepts, like crime, illness, and intelligence, can unconsciously change their behaviors<sup>24,42</sup>. For example, those who viewed local crime as a “virus” infecting their city were more likely to empathize with its perpetrators than those who viewed crime as a “beast” preying on their city; framing crime as a virus caused people to endorse rehabilitative policies over punitive measures<sup>43,44</sup>.

Because they help people make sense of unfamiliar concepts, metaphors have long been used to help people understand new technologies. Early metaphors of the Internet as a “superhighway” that could connect users to diverse digital destinations<sup>45</sup> and the later metaphor of “surfing the web” that suggests the Internet as a vehicle for exploration<sup>46</sup>. Examining the metaphors people use to describe novel developments, like AI, through analogies to familiar concepts (“AI is a friendly teacher” vs. “AI is a powerful

brain”) can provide insight into how different people are thinking and responding to this recent societal change<sup>24,25</sup>. Past work<sup>17</sup> conducted a survey of 727 participants and identified some of the most common roles that people attribute to AI, such as seeing them as including tools, servants, and assistants. We build on this work with a larger, more diverse sample of the US population and contribute a scalable approach for identifying these mechanisms at scale.

People’s implicit perceptions play a powerful role in shaping human interactions with technology. Even when people interact with the same system, they can interpret the qualities of the system in vastly different ways that affect their trust and engagement. Understanding these perceptual differences is vital for building and deploying AI responsibly. Because the inner workings of these systems are often “black boxes”, individuals rely on their intuitive beliefs when communicating with and through AI<sup>47</sup>. Similarly, the effects of interacting with AI depend on how people perceive and interpret their experiences with these complex socio-technical systems.

Past work has conducted extensive investigations into the different ways that people perceive AI-based technologies<sup>21,48,49</sup>. While people’s perceptions of AI vary depending on the context (i.e., a chatbot used for tutoring vs. a chatbot used as a medical scribe), three broad domains emerge as important drivers of human-AI interactions<sup>38,50,51</sup>: anthropomorphism, warmth, and competence.

Anthropomorphism refers to the process of viewing non-human entities as having humanlike characteristics, such as believing them to have cognitive abilities and the capacity for emotion<sup>52–54</sup>. Research on the computers-as-social-actors framework demonstrates that people often intuitively treat technologies as humanlike, as seen in people’s tendencies to apply politeness norms to conversations with chatbots without realizing it<sup>55</sup>. Indeed, numerous studies have found that people commonly anthropomorphize LLM-based technologies<sup>54,56</sup>. While increased anthropomorphism is associated with a greater willingness to trust in AI, this introduces both opportunities to support people in obtaining benefits from engaging with AI but also risks for over-reliance<sup>57,58</sup>. Prior work has also shown that users are generally skeptical of AI occupying socially intimate roles (e.g., as friends or partners), but such attitudes are often mediated by technological familiarity and acceptance<sup>59</sup>, suggesting these views may have shifted alongside recent AI advancements.

When people anthropomorphize technologies, they also view them in terms of their *warmth and competence* - two foundational dimensions of how people form impressions of social actors<sup>31,32</sup>. Warmth refers to the extent that an entity is seen as friendly, trustworthy, kind, and caring. Competence refers to an entity’s ability to act and captures beliefs about whether they are capable and intelligent. The extent to which people believe that others have warm, positive intentions and are capable of carrying out given tasks is a strong predictor of people’s attitudes, such as their trust in new technologies and willingness to adopt them in the future<sup>37–39</sup>. These dimensions are not only applied to humans<sup>60</sup> but also extend to non-human entities, including robots and other technologies<sup>32</sup>, and impact attitudes towards technology, e.g., people prefer AI agents that seem less competent<sup>37</sup>, and people rely more on AI that they perceive as warm<sup>40</sup>.

A broad literature has examined factors influencing trust in and adoption of AI, such as demographic characteristics, political orientation, and media exposure<sup>36</sup>. We focus on two core aspects of public attitudes toward AI: trust in AI and willingness to adopt AI. A certain level of trust is necessary for effective and beneficial integration of AI into society<sup>61</sup>; lack of trust and unwillingness to adopt AI may lead to under-use and missed opportunities to improve people’s lives<sup>62</sup>. However, excessive trust or premature adoption can also be dangerous, leading to over-reliance, inaccurate expectations of AI capabilities, and other harms<sup>63,64</sup>. Over-trust can also obscure the real, ongoing harms that AI systems already impose, including the amplification of social inequalities<sup>65,66</sup>, surveillance and erosion of privacy<sup>67</sup>, and the marginalization of vulnerable groups<sup>68</sup>. Moreover, public fear or backlash fueled by misplaced trust can divert attention away from addressing these concrete harms, instead focusing on speculative or

exaggerated risks<sup>69</sup>. Understanding these variables is thus critical to shaping societal outcomes of AI.

Building on prior findings that perceptions of warmth and competence predict trust and willingness to adopt AI<sup>37,38</sup>, we hypothesize that metaphors and implicit perceptions offer additional explanatory power in understanding trust and adoption. We thus examine whether these metaphorical and perceptual features can explain more variance in trust and adoption than demographics alone. These results suggest that metaphors and social perceptions can advance our understanding of how the public evaluates and engages with AI.

## Methods

**Participants & dataset.** We recruited 12,933 participants from the crowdsourcing platform Prolific between May 2023 and August 2024, with approximately 1000 individuals recruited each month, as part of a larger project understanding the US population's experiences with AI. We note that over the 16-month period, we were unable to collect data during July 2023 and May–July 2024 due to technical issues; thus, we have responses over 12 months total. We excluded a total of 30 participants for providing an invalid age (e.g., below 18, over 100;  $n = 9$ ) and for completing the survey in an unusually short time (e.g., under three minutes;  $n = 21$ ). All survey procedures were approved by the Salus Institutional Review Board, and all participants were financially compensated. We assessed the degree to which the samples each month were nationally representative of the US population by using the American National Election Studies' raking algorithms<sup>70,71</sup>. The analysis revealed that our data were nationally representative of the US population with respect to gender, ethnicity, education, and age overall, as none of the variables differed by a margin of more than 0.5%, therefore, weights were not applied. The weighting benchmarks for ensuring that our sample is representative were drawn from<sup>72</sup>'s comprehensive framework based on nine population datasets of the United States, and updated according to Pew Research's 2020 study of the American Electorate<sup>73</sup>.

The full participant demographics are as follows: 46% were men ( $n = 5929$ ), 52.3% were women ( $n = 6748$ ), and 1.8% were non-binary ( $n = 226$ ). On average, our sample was relatively diverse: 66.2% identified as white ( $n = 8536$ ), 14.2% as Black or African American ( $n = 1838$ ), 8.8% as Asian or Asian American ( $n = 1135$ ), 7.4% as Hispanic or Latino ( $n = 958$ ), 0.6% as American Indian or Native American ( $n = 79$ ), 0.2% as Pacific Islander or Alaska Native ( $n = 30$ ), and 2.5% as another race or ethnicity ( $n = 327$ ). They represented a broad range of ages, ranging from 18 to 98, with an average age of 37.6 years old (SD = 11 years). The sample was relatively well-educated: approximately 15.4% attended some college ( $n = 1992$ ), 9% received their Associate's degree ( $n = 1177$ ), 44.2% received their Bachelor's degree ( $n = 5705$ ), 18.3% received their Master's degree ( $n = 2359$ ), 2.5% received their Doctorate ( $n = 324$ ) and 2.6% completed a professional school ( $n = 340$ ), 7.5% completed high school ( $n = 964$ ) and 0.3% completed some high school ( $n = 42$ ).

After obtaining informed consent from participants, we elicited metaphors from participants by asking them: "Some people use metaphors to describe abstract concepts, like AI. What is the best metaphor for how AI works?" They were encouraged to use their own words and were assured that there were no wrong answers. We specified that the task was not about assessing their comprehension of the details of how AI works, but rather about understanding their thoughts ("This is not about accuracy as much as understanding how you are thinking"). Participants entered their answer into an open text box in the form "AI is like \_\_\_\_\_ because \_\_\_\_\_." Note that although our prompt is a simile, we use the term metaphor throughout to tie into the broader literature on conceptual metaphors and framing, which encompasses both similes and metaphors.

Next, participants completed a series of survey measures related to their experiences with and attitudes toward AI. We assessed the *frequency of each individual's AI use* by asking them to indicate if they had heard of or used 8 of the most commonly used, consumer-facing tools: ChatGPT,

DALL-E, Claude, Bard, Anthropic, Midjourney, Gemini, and Perplexity. They could also write in the names of additional AI tools that they used. To assess general trust in AI, we used a 3-item survey measure adapted from past research investigating trust in human- vs. AI-generated content<sup>74</sup>. This measure builds on the Organizational Model of Trust, which has been widely used to study trust in technologies and products<sup>75</sup> and conceptualizes trustworthiness in terms of three dimensions: (1) *ability* to achieve core tasks effectively ("I believe that AI will produce output that is accurate."), (2) *benevolence* to produce positive outcomes for the public ("I believe that AI will have a positive effect on most people."), and (3) *integrity* in facilitating user safety ("I believe that AI will be safe and secure to use). In line with past work<sup>34,74</sup>, we combined these three items by calculating their mean. The trust measure demonstrated good internal reliability (*Cronbach's*  $\alpha = 0.83$ ). Another important indicator of trust of AI focuses on assessing individuals' willingness to complete actions that would involve giving information to an AI system<sup>76</sup>. We adapted a series of 5 items that have been validated in prior international research<sup>76</sup> to examine how willing participants were to rely on information provided by AI, depend on decisions made by AI, share relevant information about themselves to enable AI systems to do tasks for them, allow their data to be used by AI, and share their feelings with AI. Items were scored on a 7-point Likert scale (1 = completely unwilling, 7 = completely willing), and composited into averages in line with past work<sup>14,76</sup>. The measure had good reliability (*Cronbach's*  $\alpha = 0.80$ ). Specifically, to measure *trust*, the survey items assessed people's perceptions of (1) AI's ability to achieve core tasks effectively ("I believe that AI will produce output that is accurate"), (2) AI's benevolence to produce positive outcomes for the public ("I believe that AI will have a positive effect on most people"), and (3) AI's integrity in facilitating user safety ("I believe that AI will be safe and secure to use."). All items were scaled on a 5-point Likert scale (1 = strongly disagree, 5 = strongly agree) and composited into a trust index in line with Ma et al. (2017) and ref. 74. For *willingness to adopt AI*, participants were asked to indicate how willing they would be to rely on information provided by AI, depend on decisions made by AI, share relevant information about themselves to enable AI systems to do tasks for them, allow their data to be used by AI, and share their feelings with AI. Items were scored on a 7-point Likert scale (1 = completely unwilling, 7 = completely willing), and composited into averages in line with past work by ref. 14.

**Ensuring data quality.** We took precautions to account for the possibility that people may use AI to generate inauthentic responses to our free-response questions by disabling copy/paste and including attention-check questions<sup>77</sup>. However, it is still possible that participants may have used AI to generate their responses to our free-response question. Given recent research indicating that AI-generated responses to survey questions tend to be systematically biased towards being more positive and similar to one another<sup>78</sup>, we chose to exclude responses that were extremely similar to those generated by AI to avoid incorporating this skew into our results. This step is important to ensuring a high-quality dataset given that previous research indicates that some participants on crowdworking platforms may be providing low-quality or inattentive responses<sup>79,80</sup>. Although we note it may be possible for participants to prompt an LLM to provide metaphors and choose the one that best aligns with their own perceptions, this metaphor-generation process would be substantively different than the imaginative, intuitive process of thinking about and eliciting metaphors of AI on their own<sup>81</sup>.

Given these considerations, we focused on excluding responses that were highly similar to outputs from ChatGPT. We evaluated the semantic similarity of participant responses with those from ChatGPT by entering the prompt into ChatGPT3.5 100 times, turning all participant-generated and GPT-generated metaphors into vectors, and using cosine similarity analysis to identify and exclude 259 metaphors that were highly similar to GPT's responses (cosine similarity > 0.85). We chose to focus on ChatGPT3.5 because it was the most common LLM accessible to respondents during the course of our data collection period. Specifically, to filter out potentially AI-generated responses, we first queried ChatGPT 100 times using variations of

**Table 1 | Examples of identifying AI-generated responses**

ChatGPT- generated response	Participant response	Cosine similarity of participant response to ChatGPT's response
AI is a sparkling symphony, each algorithmic note resonating together in perfect harmony, creating a technologically driven orchestra of intelligence.	AI is like a cosmic symphony, its algorithms echoing celestial harmonies. It's an ethereal maestro orchestrating data's celestial dance, transforming raw notes into intricate melodies of insight. With each note played, AI unravels the universe's secrets, painting a portrait of knowledge, a cosmic symphony of intelligence.	0.87
AI is the conductor of a grand symphony, seamlessly orchestrating and harmonizing the diverse instruments of information and knowledge to create a complex and awe-inspiring melody.	I imagine AI as a grand symphony orchestra conductor. Just as a conductor guides and coordinates a vast array of talented musicians to create a harmonious and captivating performance, AI orchestrates a diverse ensemble of algorithms, data, and computational power to accomplish remarkable tasks.	0.85
AI is the conductor of a grand symphony, seamlessly orchestrating and harmonizing the diverse instruments of information and knowledge to create a complex and awe-inspiring melody.	AI can be likened to a vast orchestra of interconnected musical instruments, where each instrument represents a unique algorithm or data processing unit. The conductor, symbolic of the AI's programming, skillfully orchestrates these instruments to create harmonious melodies of information, adapting and evolving its symphony based on the data it receives, akin to a never-ending, self-improving musical composition.	0.90
AI is a sparkling symphony, each algorithmic note resonating together in perfect harmony, creating a technologically driven orchestra of intelligence.	Imagine AI as a symphony of cosmic instruments. Algorithms conduct, data fuels the melodies, and users are the audience. Just as a conductor shapes music, algorithms shape AI's output. The orchestra learns and evolves, creating harmonies from the vast cosmos of information, all without consciousness but with a brilliance all its own.	0.85
AI is like a digital brain, constantly learning and evolving, capable of processing and making sense of vast amounts of information in the blink of an eye.	AI is a gardener, tending a garden of human info. The soil is data, the tools are algorithms, and the seeds of information get planted. The growing plants are the evolving insights.	0.90
AI is like a garden that continually grows and evolves, nourishing ideas and solutions, sprouting new opportunities, and fostering an environment of innovation.	AI is like a digital gardener tending to the virtual soil of data. It sows seeds of algorithms and nurtures them with data rain, allowing the garden to flourish and produce the fruits of insights and solutions, just as a real gardener cultivates a bountiful harvest	0.87
AI is like a garden that continually grows and evolves, nourishing ideas and solutions, sprouting new opportunities, and fostering an environment of innovation.	AI is a virtual garden where each and every interaction with users, plants a seed of data. Through continuous learning and cross-pollination of ideas, the garden grows and evolves into a flourishing ecosystem of knowledge, with AI acting as the wise gardener in charge of tending to the growth of understanding.	0.87

The metaphor prompt was entered into ChatGPT3.5 100 times, produce outputs such as those in the "ChatGPT-generated response" column. Next, we turned all participant-generated and GPT-generated metaphors into vectors, and used cosine similarity analysis to identify and exclude 259 metaphors that were highly similar to GPT's responses (cosine similarity > 0.85).

the prompt we gave to the participants, thus obtaining 100 responses  $m_i^{GPT}$  where  $i = 1, \dots, 100$ . For each GPT-generated response, we used the sentence embedding model *all-mpnet-base-v2* to generate a 768-dimensional embedding  $e(m_i^{GPT})$ , thus constructing a set of embeddings

$$E^{GPT} = \{e(m_i^{GPT}) | i = 1, \dots, 100\}$$

representing AI-generated metaphors. We also computed an embedding  $e(m_p)$  for each participant-written metaphor  $m_p$ . Then, for each participant-written metaphor, we measure the cosine similarity  $sim(e(m_p), e(m_i^{GPT}))$  between  $e(m_p)$  and each embedding  $e(m_i^{GPT}) \in E^{GPT}$ . If any cosine similarity was greater than 0.85 (a threshold chosen based on manual inspection of how semantically similar the metaphors are at different thresholds), we identified it as highly similar to a GPT-generated response and thus excluded it from the dataset.

We exclude 276 such participant responses, resulting in 11,790 valid metaphors. As can be seen in Table 1, this set of responses differed from the human-generated responses because they tended to be substantially longer on average, frequently used certain words and phrases (e.g., see ref. 78), and often had more elaborate language (i.e., "AI unravels the universe's secrets, painting a portrait of knowledge, a cosmic symphony of intelligence"). We note that these responses constitute < 2% of our sample and that our results - which draw upon large-scale, quantitative analysis - remain substantively unchanged with or without these examples."

**Identifying dominant metaphors through topic modeling.** To identify key thematic clusters in participants' responses, we use a quantitative, LM-based topic modeling approach. As shown in Supplementary Fig. 1, this involved (1) automatic clustering, (2) manual refinement of clusters, and (3) outlier analysis. First, we removed frequently occurring words from the response prompt (e.g., "AI" and "artificial intelligence") from all metaphors to ensure that the identified clusters were based on meaning rather than differences in phrasing. The automatic clustering approach used the state-of-the-art sentence embedding model *all-mpnet-base-v2*<sup>82</sup> to generate a 768-dimensional embedding  $e(m_p)$  that captured the meaning of each participant's metaphor of AI. Following standard practice<sup>83</sup> and based on qualitative examination of the clusters generated from varying dimensionalities, these representations were reduced to 20 dimensions using UMAP<sup>84</sup> and normalized using the L2 norm to reduce complexity prior to clustering. We then clustered these representations with HDBScan, a clustering algorithm that groups similar metaphors without forcing every metaphor into a group or pre-defining a number of groups<sup>85</sup>. The clustering was performed using the BERTopic Python package<sup>86</sup>. This process categorized about 80% of the metaphors in one of 50 clusters, while the remaining 20% were considered outliers.

Using iterative discussions and consensus-building methods that are standard in qualitative coding approaches to developing grounded theory<sup>87,88</sup>, the research team iteratively grouped the clusters based on thematic and conceptual similarities (e.g., combining a cluster about "stringing together ideas" and a cluster about "a chef combining ingredients" into the

dominant metaphor of “synthesizer”), focusing on distinctiveness between and coherence within clusters, until we reached consensus. Specifically, following the principles of Consensual Qualitative Research (CQR), each cluster was first coded independently by one of three experts by grouping it conceptually. Then, the groupings were discussed in weekly consensus meetings until unanimous agreement was reached. This dialogic adjudication balances methodological rigor with reflexivity and has been shown to yield conceptually richer results than purely reliability-based reconciliation strategies<sup>89</sup>. This resulted in a final set of 20 clusters (Note that we did not pre-constrain the number of clusters in the final set.) We refer to these 20 most frequently occurring metaphor clusters.

Finally, we re-assigned outlier metaphors that were not initially automatically assigned to a cluster as follows: For each dominant metaphor, we computed the centroid as the average of all the embeddings of the metaphors belonging to that cluster. Then, for each outlier metaphor  $m_p$ , we measured the cosine similarity of its embedding  $e(m_p)$  to each centroid.  $m_p$  was then categorized under the dominant metaphor with the highest cosine similarity if the similarity was greater than 0.6, a threshold we determined based on manual inspection of how conceptually similar the metaphors are at different thresholds. This process enabled us to automatically categorize metaphors that aligned with the broader themes identified in the manual refinement phase, but may have been too semantically different to be initially identified. This process results in 10,629 of the 11,789 metaphors being assigned to a dominant metaphor.

**Measuring implicit perceptions of AI from metaphors.** To measure implicit perceptions, we developed a scalable, systematic framework to score metaphors on the dimensions of anthropomorphism, warmth, and competence. Our approach is based on methods to automatically score any open-ended text on these dimensions, and thus, we apply them to every individual metaphor. For each dominant metaphor, we compute the mean anthropomorphism, warmth, and competence score across the individual metaphors in that cluster.

*Anthropomorphism.* To assess the extent to which people’s metaphors of AI ascribed human-like characteristics to AI, we adapted AnthroScore<sup>29</sup>, an automatic metric of implicit anthropomorphism in language. In their work, they use the masked language model RoBERTa to calculate the relative probability that a given entity (e.g., “AI”) in a text would be replaced by human pronouns versus non-human pronouns. Specifically, the degree of anthropomorphism for entity  $x$  in sentence  $s$  is measured as:

$$A(s_x) = \log \frac{P_{HUMAN}(s_x)}{P_{NON-HUMAN}(s_x)},$$

where

$$P_{HUMAN}(s_x) = \sum_{w \text{ human pronouns}} P(w), P_{NON-HUMAN}(s_x) = \sum_{w \text{ non-human pronouns}} P(w),$$

where  $P(w)$  is the model’s outputted probability of replacing the mask with the word  $w$ . We apply AnthroScore to the entities  $x \in \{\text{it, AI, artificial intelligence}\}$  to measure the anthropomorphism of AI in each metaphor  $m_p$ . If the metaphor does not contain any of these entities, we use the spacy package<sup>61</sup> to identify whether the metaphor is only a noun phrase. If so, we prepend the phrase “AI is” and then apply AnthroScore to the now-present term “AI” to measure anthropomorphism. Since this score is a relative log probability, a score greater than 0 suggests that the entity is more likely to be human, and a score less than 0 suggests that the entity is more likely to be non-human. For example, in the metaphor “AI is a teacher” (and the entity “AI”), we compute the probability of the sentence “He is a teacher” and the sentence “It is a teacher.” Because the former is a much more probable sentence, this metaphor has AnthroScore  $> 0$ , indicating that it anthropomorphizes AI. For interpretability, we use an indicator score of

whether  $A(s_x)$  is greater than 0:

$$A^{bin}(m_p) = \mathbb{1}(A(m_p) > 0)$$

This binary form allows us to easily and directly measure the percentage of metaphors that are anthropomorphic (i.e.,  $A^{bin}(m_p) = 1$ ), as reported in the following sections.

**Measuring warmth and competence.** We were also interested in understanding the extent to which people’s metaphors characterized AI as being warm and competent - two fundamental dimensions of social perception that are well-established psychological constructs for estimating perceptions of people<sup>60</sup> and anthropomorphized non-human entities<sup>32</sup>. These dimensions reflect, respectively, perceived friendliness, trustworthiness, and empathy (warmth) versus capability, intelligence, and effectiveness (competence).

To assess these qualities in the context of AI metaphors, we applied the **contextualized semantic axes method**<sup>90,91</sup>. This approach quantifies conceptual dimensions, such as warmth and competence, by measuring the semantic similarity between a given text (in our case, a metaphor about AI) and the constructed contextualized semantic axis. This method enables us to position words or phrases on a continuous spectrum between two extremes (e.g. warm vs. cold or competent vs. incompetent) using contextualized embeddings derived from a sentence embedding model, specifically *all-mpnet-base-v2*.

Following the method of ref. 30 for constructing semantic axes for warmth and competence, we employed **validated lexicons** for warmth and competence as anchors to define these axes. These lexicons include words or phrases strongly associated with high or low warmth and competence, enabling robust alignment of the embedding space with human perceptions. For each metaphor, we computed its cosine similarity to the axis, resulting in a score (between -1 and 1) that indicates its warmth and competence. By assigning a score between -1 and 1 for warmth and competence for each metaphor, this enables the systematic analysis of people’s perceptions of AI. Specifically, the semantic axes are defined as follows: for a given dimension  $D$  (warmth or competence), we construct the semantic axis by taking the mean of the embeddings for each word in the lexicon that are positively associated with that dimension, and subtract the mean of the embeddings for each word in the lexicon that are negatively associated with it:

$$A_D = \frac{1}{k} \sum_{i=1}^k e(w_i) - \frac{1}{m} \sum_{j=1}^m e(w'_j)$$

where  $w_i/w'_j$  is a word in the set of words positively/negatively associated with that dimension respectively. We again use *all-mpnet-base-v2* to compute the embeddings. For each metaphor  $m_p$  embedded as  $e(m_p)$ , we compute its warmth and competence as the cosine similarity of the embedding  $e(m_p)$  to this axis, i.e.,

$$\text{Warmth}(m_p) = \cos(e(m_p), A_{\text{warmth}}); \text{Competence}(m_p) = \cos(e(m_p), A_{\text{competence}}),$$

resulting in two scores (each between -1 and 1) that represent the metaphor’s warmth and competence respectively. Examples of metaphors with varying warmth and competence scores are in Table 2.

**Multivariable regression.** Next, we measured how dominant metaphors and implicit perceptions help explain two attitudinal variables with consequential outcomes: trust in AI and willingness to adopt AI. We conducted a pair of hierarchical multiple regressions that entered (1) participant demographics, (2) average AI use, (3) the number of AI tools heard of, (4) dominant metaphors, and (5) implicit perceptions to examine the extent to which each block explained additional variance in trust in AI and willingness to adopt AI. This sequential approach, which has been previously used to understand the explanatory power of

**Table 2 | Metaphors with varying levels of warmth and competence**

<p>Positive warmth, negative competence:                  "It's like a dog. On one hand, it can be gentle and loving. However, you don't know if it may suddenly bite."                  "Like an imaginary friend"                  "AI is like Your uncle is who right 50% of time. because They have a lot to say but don't understand the context."                  "AI is like having a forever friend because they cant get up and leave"</p>	<p>Positive warmth, positive competence:                  "AI is like having a 'phone a friend' lifeline on the TV gameshow, 'Who wants to be a millionaire.'                  You can ask for assistance from a more knowledgeable entity for personal assistance or gain."                  "AI is like a wise old friend because it knows a lot and can give very helpful advice."                  "It makes me think of a fairy god mother. All knowing and there to help you out"                  "It's like a robot that can learn over time"</p>
<p>Negative warmth, negative competence:                  "It grinds my gears"                  "Like a turtle trying to run like a rabbit."                  "Its a computer predicting the next word It has no understanding"                  "Shoddy thief"</p>	<p>Negative warmth, positive competence:                  "A detective that collects clues from different recourses, analyzes them and uses that information to make a prediction or solve a problem"                  "Its a computer system that is loaded with a bunch of data. ALL the data is tied to keywords so when the key word us used a bunch of data with that tag is pulled up."                  "It's like a copy machine that takes info and tries to replicate it"                  "Putting human knowledge into a system and then having it spit out I got stood at you. It could be in the form or writing or even art."</p>

independent variables (e.g. refs. 92,93), allowed us to better understand the additional explanatory value provided by considering metaphors and their implicit perceptions beyond assessments of AI use and demographic factors.

All predictors were standardized to the range [0,1] to ensure comparability of effect sizes. We first conducted a regression to explain the dependent variable (trust or adoption) using a foundational set of predictors related to demographics and familiarity with AI: time, frequency of AI tool use, number of AI tools heard of, gender, race/ethnicity, age, education level, and work level. (To avoid collinearity, we drop the most common value for the categorical variables of gender, race/ethnicity, and education level. That is, the regression is relative to the variables of Male gender, white race/ethnicity, and Bachelor's degree education level. In the second step, we introduce the second block of variables, which consists of 20 variables, each representing the presence of a dominant metaphor (tool, robot, assistant, etc.). Finally, we incorporated the third block of variables, which are the implicit perceptions measured by our framework: anthropomorphism, warmth, and competence. Model assumptions, including linearity, multicollinearity, and homoscedasticity, were also checked to ensure the validity of the results. Our experiments were not preregistered.

## Results

### Dominant conceptualizations of AI: Most people understand AI through one of 20 dominant metaphors

As detailed in Methods, we used a combination of automatic clustering and manual iterative coding to identify the 20 dominant clusters of metaphors that people used to conceptualize AI in our dataset (Table 3). The most prevalent framed AI as being like a tool, such as a calculator or Swiss Army knife (10%); an external biological brain capable of reasoning and logic (10%); and a powerful search engine capable of navigating large databases (9%). However, the metaphors were still highly diverse. Some viewed AI as an intelligent teacher ("AI is like a professor because it always has the answer", 3%) whereas others viewed it as akin to a child ("It needs to learn, but is happy to provide output and is proud of it, 4%). People also drew inspiration from popular conceptions of AI, likening it to the Terminator and other humanoid robots (8%). In contrast to its mechanical nature, many also likened AI to pets ("It's like training your dog to do something", 2%) or unexplored wilderness ("AI is like the ocean, there is so much uncertainty and so much to discover", 2%).

### Implicit perceptions of AI: Metaphors reveal temporal shifts in people's implicit perceptions of AI

The metaphors people use to describe AI can reveal the implicit perceptions they hold about AI. Applying our framework to measure implicit perceptions from the metaphors, we compute mean anthropomorphism, warmth, and competence scores for each dominant metaphor. This revealed that people hold nuanced views of how human-like they perceive AI to be, and generally view AI to be both warm and competent (Fig. 1). For example,

describing AI as being like "a child" humanizes it while highlighting its potential to learn. Seeing AI as a "Swiss Army knife" frames it as inherently mechanical and capable of many functions. A notable exception is the metaphor of AI as a "thief", which reflects a negative perception of its intentions.

Analyzing month-over-month shifts revealed how public perceptions of AI have changed over time (Fig. 2). People anthropomorphized AI more over time (+34% over the 12 months,  $r^{10} = 0.80$ ,  $p = 0.002$ , 95% CI [0.42, 0.94]), becoming more likely to describe AI as being like a teacher ( $r^{10} = 0.75$ ,  $p = 0.005$ , 95% CI [0.30, 0.92]), a friend ( $r^{10} = 0.75$ ,  $p = 0.005$ , 95% CI [0.32, 0.93]), or an assistant ( $r^{10} = 0.60$ ,  $p = 0.040$ , 95% CI [0.04, 0.87]). Conversely, people became less likely to see AI as a distinctly non-human entity such as a computer ( $r^{10} = -0.92$ ,  $p < 0.001$ , 95% CI [-0.98, -0.74]), a search engine ( $r^{10} = -0.63$ ,  $p = 0.029$ , 95% CI [-0.88, -0.08]), a mirror ( $r^{10} = -0.70$ ,  $p = 0.012$ , 95% CI [-0.91, -0.20]), or a synthesizer ( $r^{10} = -0.91$ ,  $p < 0.001$ , 95% CI [-0.98, -0.71]).

People also saw AI as being warmer over time (+41% over 12 months,  $r^{10} = 0.62$ ,  $p = 0.030$ , 95% CI [0.08, 0.88]). Notably, however, this did not correspond to an increased perception that AI is competent; implicit perceptions of AI as competent instead decreased by 8% over time ( $r^{10} = -0.60$ ,  $p = 0.039$ , 95% CI [-0.87, -0.04]). Accordingly, the highest-warmth-dominant metaphors (friend, assistant, and god (god:  $r^{10} = 0.64$ ,  $p = 0.025$ , 95% CI [0.11, 0.89])) became more common over time, whereas several of the highest-competence metaphors (computer, synthesizer, and search engine) became less common.

### Dominant metaphors and implicit perceptions help explain trust and willingness to adopt AI

Understanding how people think and feel about AI can help explain how people respond to the integration of AI in society. We conducted a pair of hierarchical multiple regression analyses to examine the extent to which dominant metaphors and perceptions can explain variance in 1) people's trust in AI and 2) their willingness to adopt AI, relative to demographic and AI familiarity differences, over and above demographic differences and individuals' experiences with AI (Fig. 3).

For trust, the baseline model including demographics and AI familiarity was significant,  $F(14, 9480) = 152.39$ ,  $p < 0.001$ ,  $R^2 = 0.184$ , adj.  $R^2 = 0.182$ . Adding dominant metaphors significantly improved model fit,  $F(32, 9462) = 77.20$ ,  $p < 0.001$ ,  $R^2 = 0.207$ , adj.  $R^2 = 0.204$ ,  $\Delta R^2 = 0.023$ ,  $F\text{-change}(18, 9462) = 15.47$ ,  $p < 0.001$ . Including implicit perceptions (anthropomorphism, warmth, competence) further improved the model,  $F(35, 9459) = 96.90$ ,  $p < 0.001$ ,  $R^2 = 0.264$ , adj.  $R^2 = 0.261$ ,  $\Delta R^2 = 0.057$ ,  $F\text{-change}(3, 9459) = 243.60$ ,  $p < 0.001$ .

For willingness to adopt AI, the baseline model (demographics + AI familiarity) explained a smaller but still significant portion of variance,  $F(14, 9480) = 117.92$ ,  $p < 0.001$ ,  $R^2 = 0.148$ , adj.  $R^2 = 0.147$ . Adding metaphors increased explanatory power,  $F(32, 9462) = 59.19$ ,  $p < 0.001$ ,  $R^2 = 0.167$ , adj.  $R^2 = 0.164$ ,  $\Delta R^2 = 0.018$ ,  $F\text{-change}(18, 9462) = 11.65$ ,  $p < 0.001$ . Finally, adding perceptions (anthropomorphism, warmth, competence) yielded a

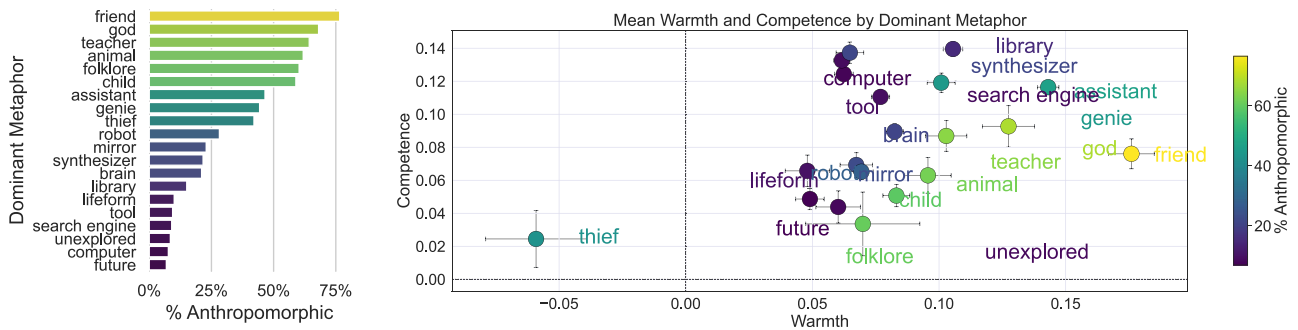
**Table 3 | Dominant metaphors used to describe AI**

Metaphor	Description	Example	Freq. (%)	Pearson corr. over time	Anthropomorphism (M)	Warmth (M)	Competence (M)
Tool	AI is seen as similar to technological tools like engines, calculators, and appliances and analog tools, like hammers and Swiss Army knives	"AI is like a very smart scientific calculator with access to the internet. It still needs input to have an accurate output. Just like how the higher end calculators work, the output may be in the wrong format and need to be adjusted in order for the data to be used." "AI works like the engine on a car, the more you put into it the more you get out of it"	10%	0.91 (ns)	-3.39	0.08	0.11
Biological brain	AI is seen as a powerful brain, capable of human-level thinking, intelligence, and reasoning but detached from other human-like qualities	"AI is like an external brain you can access to help you solve problems" "AI is like a human brain but without all the information that is unnecessary or distracting. It also operates without emotions."	10%	-0.48 (ns)	-2.08	0.08	0.09
Search engine	AI is seen as a tool that can sift through data and information on the internet, like a search engine, directory, or database navigator	"AI works similarly to Google in the sense that you can type in a question, and find your answer with the results. However, with AI, they have an entire interface that allows the AI to find what it thinks is the best answer, instead of having to research through several options." "AI is kind of similar to the existence of a 411 like telephone hotline. You can ask it anything and they are able to direct you to the correct resource if not answer it themselves."	9%	-0.61 (p = 0.029)	-3.40	0.06	0.13
Work assistant	AI is seen as an assistant or employee who can help users complete tasks and find answers	"To me, AI is like having a personal assistant. It's something that you can delegate certain tasks to be performed. But it's not always perfect and can be prone to making mistakes." "AI is like a good to average personal assistant because you can rely on it to get tasks done right... most of the time."	8%	0.63 (p = 0.040)	-0.38	0.14	0.12
Humanoid robot	AI is seen as an embodied robot that resembles humans	"AI is like robot that tries to think because it processes information but has no feelings." "AI as an embodied robot that resembles humans, as in science fiction, like the Terminator."	8%	0.7 (p = 0.012)	-1.38	0.07	0.07
Intelligent computer	AI is seen as a powerful computer program or software with enhanced data processing and analytic abilities	"AI is kind of like gathering all of the intelligence in the world together into a collective mind and then using that information to extrapolate information and make complex decisions."	7%	-0.92 (p < 0.001)	-3.40	0.06	0.12
Library	AI is seen as a source of extensive knowledge, like a library or encyclopedia	"AI is like a limitless thesaurus that can talk because it compiles information across what seems like limitless databases and compiles information, makes recommendations, and can provide feedback with direction." "AI is like a librarian that can recite information about every book on the shelves."	5%	0.62 (ns)	-2.74	0.11	0.14
The future	AI is seen as a force that will shape the future in both positive and negative ways	"A tech fad that will provide some actually useful uses in the coming years, but for now will mostly be used to advance data security violations for companies that already make too much money."	5%	0.39 (ns)	-3.13	0.05	0.05
Magic genie	AI is seen as a mystical or mysterious being, such as a genie, wizard, or fortune teller	"The best metaphor is it being a Genie as it can create positive and negative things." "Magic, in the way it's able to interpret one's words and the type of responses that it gives. It's impressive."	4%	0.46 (ns)	-0.58	0.10	0.12
Mirror	AI is seen as a mirror that reflects, imitates, or mimics humans	"It feels like a distorted mirror." "AI is like Talking to an echo chamber, because So much of what it says is based off of what people have told it and what it has read."	4%	0.69 (p = 0.012)	0.33	0.08	0.05
Child	AI is seen as a developing child that can learn and grow	"AI is like a child that needs to learn, but is happy to provide output and is proud of it." "AI works like a child. It absorbs and learns from those around it. It gathers information and eventually is able to do those things and create things on its own."	4%	0.64 (p = 0.039)	-1.92	0.07	0.07
			3%		-2.11	0.06	0.14

**Table 3 (continued) | Dominant metaphors used to describe AI**

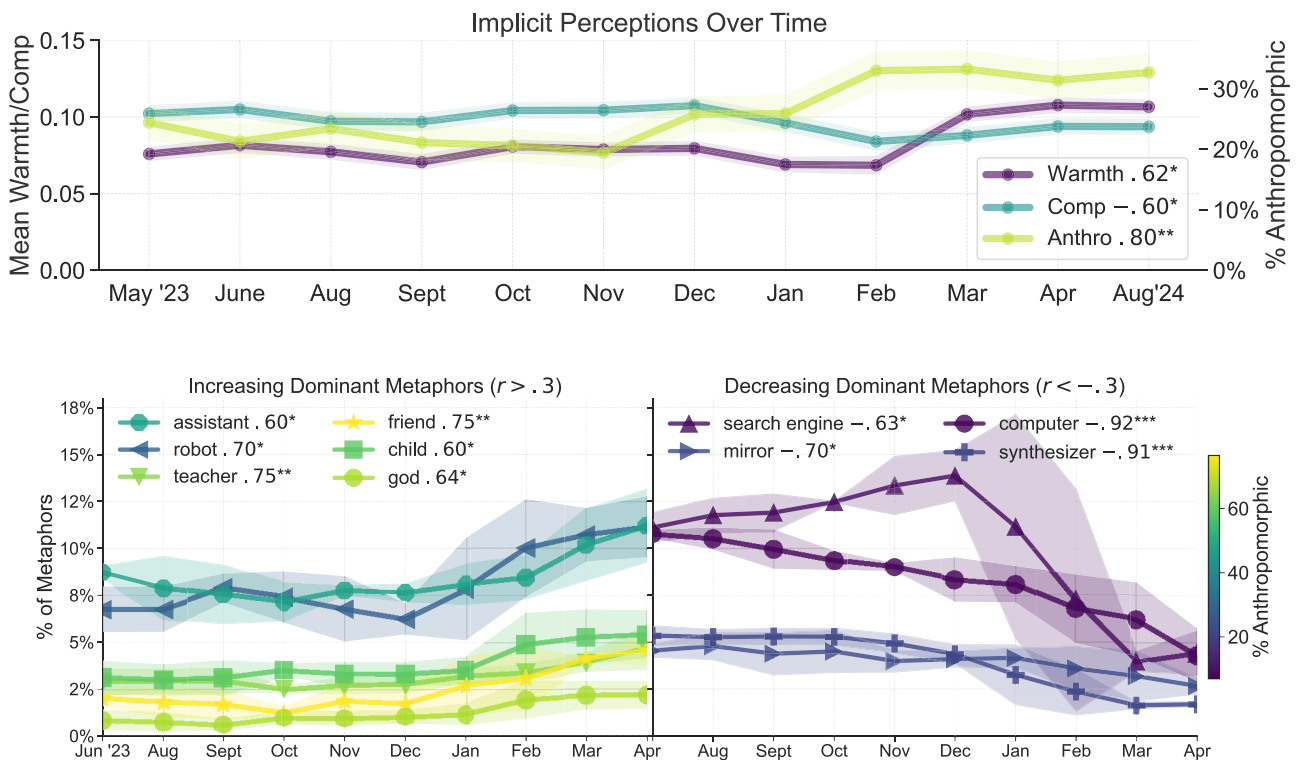
Metaphor	Description	Example	Freq. (%)	Pearson corr. over time	Anthropomorphism (M)	Warmth (M)	Competence (M)
Creative synthesizer	AI is seen as a synthesizer that is able to combine elements creatively to form something new	"AI is like a painter. You can give AI all the materials such as canvas, paint, paint brushes and a subject to draw. With these materials, AI can put those things together and create something." "AI is like an extension of my artistic abilities because through guidance I can get images close to what I would complete if I was using a camera and a subject."		-0.90 (p < 0.001)			
Teacher	AI is seen as a provider of knowledge and advice, like a teacher, professor, or mentor	"AI is like a professor because it always has the answer." "AI is like that smart kid in class that does your homework for you. It's a bit like a helper."	3%	0.75 (p = 0.005)	0.90	0.10	0.09
Friend	AI is seen as a friend with varying levels of reliability, closeness, and knowledge	"AI is like a friend you never had, but maybe never wanted." "An AI works like a genius friend who you can ask pretty much anything but who doesn't have a high emotional capacity."	2%	0.76 (p = 0.005)	1.51	0.18	0.08
Living nature	AI as wilderness that can self-propagate and grow, potentially beyond human control, such as plants and ecosystems	"AI is like a waterbead because it can take what little you put into it and expand to give you more information." "AI is like A tree of knowledge because it is continuously growing and expanding intelligence"	2%	0.25 (ns)	-2.73	0.05	0.07
Animals and pets	AI as various animals, possibly serving as companions (like pets)	"AI is like A furry that is too smart because it still seems arbitrary with some responses." "It's like how an animal learns to mimic behaviors"	2%	-0.21 (ns)	0.11	0.10	0.06
Unexplored wilderness	AI as aspects of nature that remain largely unexplored by humans, such as the ocean, jungle, and outer space	"AI is like a jungle because it is vast and has interconnected network of pathways and holds a diverse variety of life forms." "AI is like the ocean, there is so much uncertainty and so much to discover"	2%	0.49 (ns)	-2.75	0.06	0.04
Omnipresent god	AI as an always-there, godlike, omnipresent, and all-knowing presence.	"AI is like a God because it knows everything" "A deity or divine creature that has the knowledge and data to make informed decisions on a grand scale but lacks the consequences of decisions on an individual level. Basically, they are able to access more data than humans to make logical decisions but lack the empathy and morality that keep decisions fair for most."	1%	0.66 (p = 0.025)	1.01	0.13	0.09
Folklore characters	AI as miscellaneous figures from fairytales and folklore, such as fairy godmother, Icarus, angel, satan/devil	"AI is like a fairy godmother - there to help you with whatever you need" "AI is like A ghost because it's there and can help you but also might not be inherently good."	1%	0.4 (ns)	0.77	0.07	0.03
Thief	AI as a thief of and threat to others' work and livelihoods	"AI is basically plagiarism. It uses existing work and copies it." "AI is like putting documents through a shredder, then rearranging the shreds into a new 'document' because it's just taking things that already exist and stealing from just enough different sources that it's not completely recognizable."	0.5%	0.25 (ns)	-1.05	-0.06	0.02

We identify 20 dominant metaphors (ordered by frequency (%)) from the 12,000 metaphors collected using a combination of LM-based clustering and qualitative coding. *n* is the count of the dominant metaphor, and *r* is the Pearson's correlation between its frequency and time; positive/negative correlations indicate increases/decreases over time respectively. \*, \*\*, \*\*\*, and (ns) denote  $p < 0.05$ ,  $p < 0.01$ ,  $p < 0.001$ , and no statistical significance respectively.



**Fig. 1 | Mean anthropomorphism, warmth, and competence of dominant metaphors for AI (n = 11,790).** Error bars show 95% confidence intervals of the mean (see Table 1 for n of each dominant metaphor). Colors correspond to anthropomorphism rate. The dominant metaphors (n = 20) vary widely in anthropomorphism. They are all positive in both mean warmth and mean competence (except “thief”); based on qualitative inspection of the individual metaphors, we find that this is because although some of the dominant metaphors reflect concepts that may have negative connotations, participants focus on warm and competent aspects of these concepts in their responses. For example, when

describing AI as a “child,” participants often describe its learning and growing capabilities, and when describing AI as a “genie,” participants focus on its capabilities, which seem magical in their scope. For the dominant metaphor of “unexplored realm”, rather than expressing fear of the unknown, participants’ metaphors discuss the potential for discovery and capacity for “wonder and beauty.” Even for the dominant metaphor of “thief,” participants’ responses describe, for instance, that AI is stealing work and that it is able to accomplish many tasks through this stealing.



**Fig. 2 | Shifts in public perceptions of AI over time (n = 11,790).** Monthly mean warmth and competence scores (left y-axis, range [-1, 1]) and percentage of anthropomorphic metaphors (right y-axis). Shaded bands indicate 95% confidence intervals across monthly means. Month-by-month prevalence of each dominant metaphor with significant temporal change ( $|r| > 0.3, p < 0.05$ ). Lines are color-

coded by the proportion of anthropomorphic metaphors within each dominant metaphor. Shading represents a 3-month rolling average. Anthropomorphic and warm metaphors increase in frequency over time, whereas competence-focused metaphors decrease. Missing months reflect survey interruptions due to technical issues.

substantial improvement,  $F(35, 9459) = 71.49, p < 0.001, R^2 = 0.209, \text{adj. } R^2 = 0.206, \Delta R^2 = 0.042, F\text{-change}(3, 9459) = 169.04, p < 0.001$ .

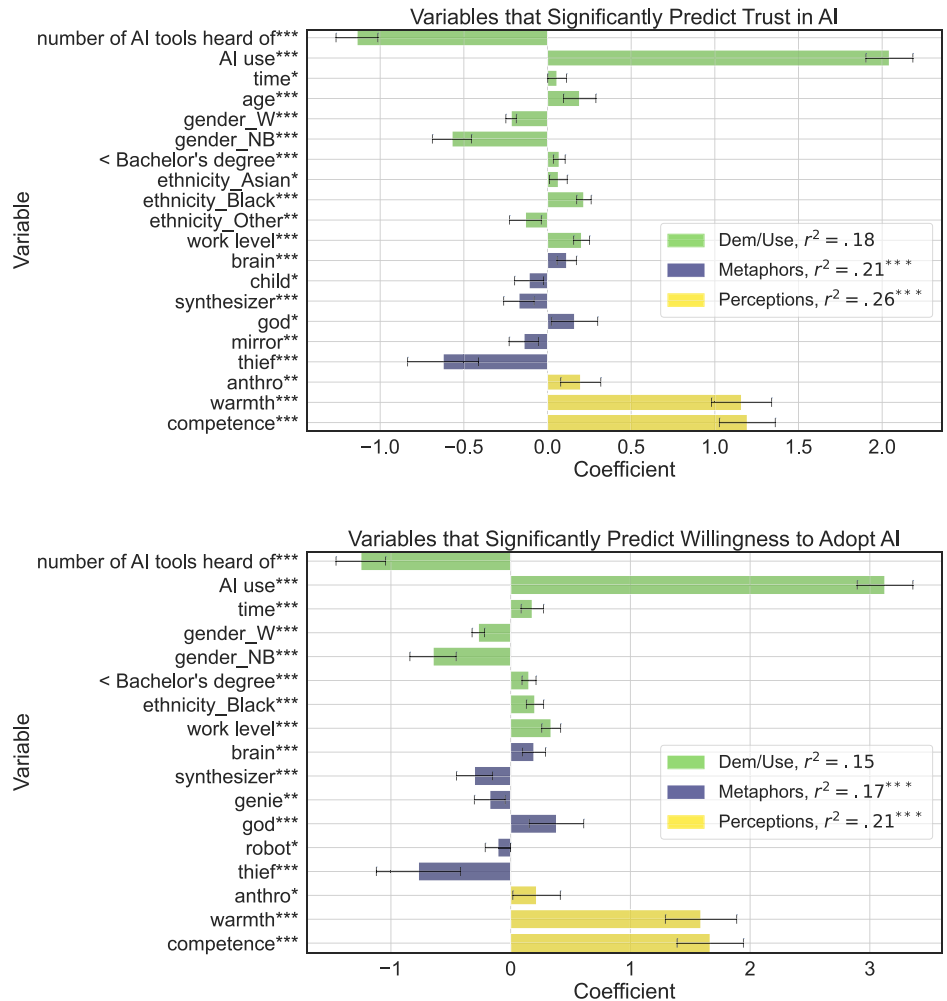
Overall, the first block of variables (demographics + AI familiarity) explained 18% and 15% of the variance in trust and willingness to adopt AI, respectively ( $\text{adj. } R^2 = 0.18, 0.15$ ). The second block (dominant metaphors) explained an additional 2–3% of variance ( $p < 0.001$ ), and the third block (implicit perceptions) explained an additional 4–6% ( $p < 0.001$ ). Together, the dominant metaphors and implicit perceptions account for substantially

more variance—44% more for trust and 40% more for willingness to adopt—than demographics and frequency of use alone (Fig. 3; see detailed results in Table 4).

**Specific metaphors and perceptions are associated with trust and adoption**

Regression coefficients from the full model (Fig. 3) show that, beyond AI use and familiarity, warmth and competence are the strongest predictors of trust

**Fig. 3 | Dominant metaphors and implicit perceptions of AI help explain trust in AI and willingness to adopt AI (n = 11,790).** Standardized regression coefficients ( $\beta$ ) from the combined model incorporating three predictor blocks: demographics + AI use (green), dominant metaphors (purple), and implicit perceptions (yellow). Each bar shows the mean standardized coefficient; error bars denote 95% confidence intervals. Asterisks indicate significance (\* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ ; two-sided t-tests on model coefficients). All predictors were normalized to [0, 1] for comparability. Frequency of AI use was the strongest predictor of both trust and adoption, followed by the number of AI tools heard of, warmth, and competence. Adding metaphors and perceptions increased explained variance by 44% and 40%, respectively, relative to demographic + use variables alone.



and AI adoption. For trust, warmth ( $b = 1.804$ ,  $SE = 0.142$ ,  $t(9467) = 12.69$ ,  $p < 0.001$ , 95% CI [1.526, 2.083]) and competence ( $b = 1.821$ ,  $SE = 0.124$ ,  $t(9467) = 14.63$ ,  $p < 0.001$ , 95% CI [1.577, 2.064]) were both highly significant predictors. For adoption, warmth ( $b = 2.455$ ,  $SE = 0.234$ ,  $t(9467) = 10.52$ ,  $p < 0.001$ , 95% CI [1.998, 2.913]) and competence ( $b = 2.499$ ,  $SE = 0.204$ ,  $t(9467) = 12.23$ ,  $p < 0.001$ , 95% CI [2.098, 2.899]) again showed strong positive effects.

This is likely due to warm and competent metaphors such as *assistant*, *friend*, *teacher*, and *library*, which stand out as predicting trust and adoption positively in a dominant-metaphor-only regression (Supplementary Fig. 2). For trust, the coefficients were: *friend* ( $b = 0.297$ ,  $SE = 0.054$ ,  $t(10616) = 5.45$ ,  $p < 0.001$ , 95% CI [0.190, 0.403]), *assistant* ( $b = 0.233$ ,  $SE = 0.034$ ,  $t(10616) = 6.87$ ,  $p < 0.001$ , 95% CI [0.167, 0.300]), *teacher* ( $b = 0.092$ ,  $SE = 0.050$ ,  $t(10616) = 1.86$ ,  $p = 0.063$ , 95% CI [-0.005, 0.190]), and *library* ( $b = 0.242$ ,  $SE = 0.040$ ,  $t(10616) = 6.10$ ,  $p < 0.001$ , 95% CI [0.164, 0.320]). For adoption, effects were similarly positive: *friend* ( $b = 0.487$ ,  $SE = 0.088$ ,  $t(10616) = 5.53$ ,  $p < 0.001$ , 95% CI [0.314, 0.659]), *assistant* ( $b = 0.368$ ,  $SE = 0.055$ ,  $t(10616) = 6.69$ ,  $p < 0.001$ , 95% CI [0.260, 0.476]), *teacher* ( $b = 0.168$ ,  $SE = 0.080$ ,  $t(10616) = 2.10$ ,  $p = 0.036$ , 95% CI [0.011, 0.326]), and *library* ( $b = 0.310$ ,  $SE = 0.064$ ,  $t(10616) = 4.83$ ,  $p < 0.001$ , 95% CI [0.184, 0.435]). These effects appear mediated by the warmth and competence variables in the full model (Fig. 3).

Anthropomorphism also predicts positive trust and adoption, but to a lesser extent. For trust, anthropomorphism ( $b = 0.011$ ,  $SE = 0.003$ ,  $t(9467) = 3.61$ ,  $p < 0.001$ , 95% CI [0.005, 0.016]) showed a small but

significant effect, and for adoption, anthropomorphism ( $b = 0.011$ ,  $SE = 0.005$ ,  $t(9467) = 2.27$ ,  $p = 0.023$ , 95% CI [0.001, 0.020]) also predicted modest increases.

Notably, even in combination with the perceptual variables, *thief* and *synthesizer* strongly predicted low trust and reluctance to adopt. For trust, *thief* ( $b = -0.656$ ,  $SE = 0.110$ ,  $t(9467) = -5.96$ ,  $p < 0.001$ , 95% CI [-0.872, -0.440]) and *synthesizer* ( $b = -0.220$ ,  $SE = 0.047$ ,  $t(9467) = -4.63$ ,  $p < 0.001$ , 95% CI [-0.313, -0.127]) were both negative predictors. For adoption, *thief* ( $b = -0.801$ ,  $SE = 0.181$ ,  $t(9467) = -4.43$ ,  $p < 0.001$ , 95% CI [-1.155, -0.446]) and *synthesizer* ( $b = -0.356$ ,  $SE = 0.078$ ,  $t(9467) = -4.57$ ,  $p < 0.001$ , 95% CI [-0.509, -0.203]) similarly reduced adoption. In contrast, *brain* and *god* were linked to higher trust and adoption: for trust, *brain* ( $b = 0.118$ ,  $SE = 0.030$ ,  $t(9467) = 3.87$ ,  $p < 0.001$ , 95% CI [0.058, 0.178]) and *god* ( $b = 0.163$ ,  $SE = 0.072$ ,  $t(9467) = 2.27$ ,  $p = 0.023$ , 95% CI [0.023, 0.303]); for adoption, *brain* ( $b = 0.204$ ,  $SE = 0.050$ ,  $t(9467) = 4.08$ ,  $p < 0.001$ , 95% CI [0.106, 0.302]) and *god* ( $b = 0.387$ ,  $SE = 0.118$ ,  $t(9467) = 3.29$ ,  $p = 0.001$ , 95% CI [0.156, 0.617]).

Building on previous work on the nuanced relationship between anthropomorphism and trust<sup>40,94</sup>, we find specific metaphors that reveal what topics are salient to people's trust and adoption: conceptualizations of AI as a helpful human-like entity ("friend", "teacher", "assistant"), or as a source of vast amounts of knowledge ("library", "god", "brain") facilitate trust and adoption. In contrast, anthropomorphic conceptualizations of AI as a human-like "thief" are often related to the widespread public concerns about copyright issues and AI plagiarizing or stealing creative work<sup>95,96</sup>.

**Table 4 | Hierarchical regression results of demographic and use variables, dominant metaphors, and implicit perceptions on trust in AI and willingness to adopt AI**

	Trust in AI			Willingness to adopt AI		
	$\beta$	adjusted $r^2$	$\Delta r^2$	$\beta$	adjusted $r^2$	$\Delta r^2$
Block 1 - Demographics and use		0.12			0.10	
frequency of AI tool use	1.54***			2.71***		
non-man	-0.22***			-0.28***		
age	0.38***			0.48***		
non-white	0.13***			0.10***		
time	0.16***			0.35***		
Block 2 - Dominant metaphors		0.15	0.03***		0.13	0.02***
assistant	0.00 (ns)			0.04 (ns)		
brain	0.15***			0.23***		
child	-0.17***			-0.06 (ns)		
synthesizer	-0.22***			-0.39***		
search engine	-0.10**			-0.15*		
folklore character	-0.05 (ns)			-0.18 (ns)		
friend	0.03 (ns)			0.11 (ns)		
future	-0.03 (ns)			-0.06 (ns)		
genie	-0.02 (ns)			-0.15*		
god	0.16*			0.40***		
library	0.04 (ns)			0.00 (ns)		
computer	-0.00 (ns)			-0.05 (ns)		
mirror	-0.17***			-0.07 (ns)		
lifeform	0.08 (ns)			-0.03 (ns)		
robot	0.02 (ns)			-0.07 (ns)		
teacher	-0.04 (ns)			-0.02 (ns)		
thief	-0.60***			-0.69***		
tool	0.01 (ns)			-0.00 (ns)		
wilderness	0.01 (ns)			-0.03 (ns)		
Block 3 - Implicit perceptions		0.21	0.06***		0.18	0.05***
anthropomorphism	0.24***			0.20*		
warmth	1.26***			1.68***		
competence	1.23***			1.72***		

\*, \*\*, \*\*\*, and (ns) denote  $p < 0.05$ ,  $p < 0.01$ ,  $p < 0.001$ , and no statistical significance respectively.

**Demographic differences reveal disparities in trust and adoption**

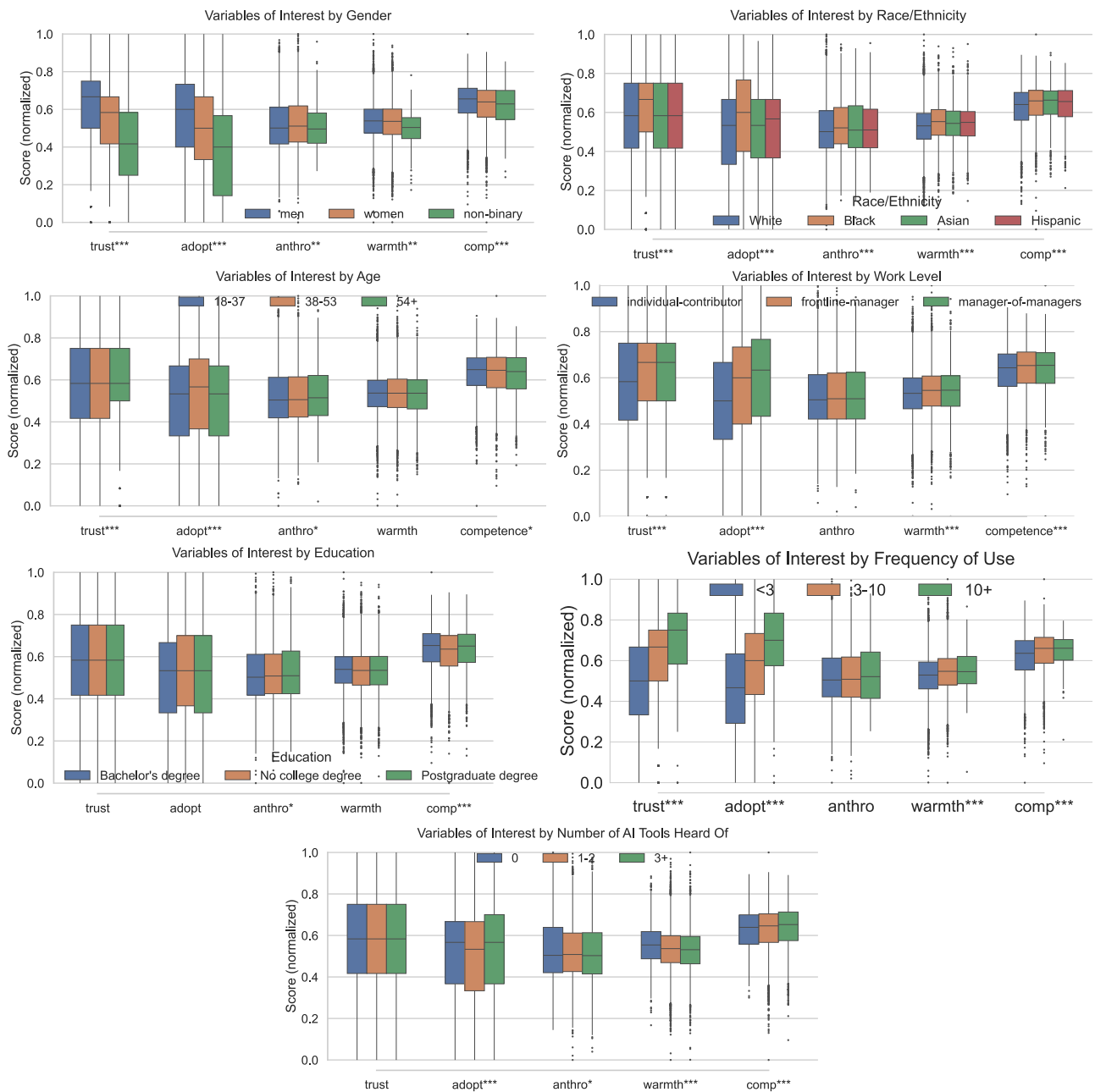
Finally, it is important to recognize that people from different backgrounds think about AI differently. Dominant metaphors and implicit perceptions provide insights into these phenomena, though we emphasize the need for additional research to establish causal relationships and a deeper understanding. Figure 4 shows differences in trust, adoption, and perceptions, and Supplementary Fig. 3 shows detailed differences in dominant metaphors for each demographic.

**Gender**

For trust, women ( $M = 0.54$  [0.54, 0.55]) and non-binary participants ( $M = 0.41$  [0.37, 0.44]) reported substantially lower scores than men ( $M = 0.62$  [0.61, 0.63]); pairwise comparisons showed that women trusted AI less than men,  $t(10175.27) = -18.98$ ,  $p < 0.001$ ,  $\Delta M = -0.08$  [-0.09, -0.07], while non-binary participants trusted AI markedly less than both women,  $t(186.24) = 7.64$ ,  $p < 0.001$ ,  $\Delta M = 0.14$  [0.10, 0.17], and men,  $t(187.68) = 12.03$ ,  $p < 0.001$ ,  $\Delta M = 0.21$  [0.18, 0.25]. A similar pattern emerged for adoption: women ( $M = 0.48$  [0.48, 0.49]) and non-binary participants ( $M = 0.39$  [0.35, 0.43]) reported lower willingness than men ( $M = 0.56$  [0.55, 0.56]); women < men,  $t(10102.52) = -16.60$ ,  $p < 0.001$ ,  $\Delta M = -0.07$  [-0.08, -0.07]; women > non-binary,

$t(185.68) = 4.89$ ,  $p < 0.001$ ,  $\Delta M = 0.10$  [0.06, 0.13]; and men > non-binary,  $t(187.48) = 8.66$ ,  $p < 0.001$ ,  $\Delta M = 0.17$  [0.13, 0.21]. Both women ( $M = 0.53$  [0.53, 0.54]) and non-binary participants ( $M = 0.50$  [0.48, 0.51]) described AI in less warm terms than men ( $M = 0.54$  [0.53, 0.54]); women < men,  $t(11362.23) = -2.44$ ,  $p = 0.015$ ,  $\Delta M = -0.01$  [-0.01, -0.00]; and non-binary < men,  $t(207.29) = 5.98$ ,  $p < 0.001$ ,  $\Delta M = 0.04$  [0.03, 0.06]. A comparable pattern held for competence, with women ( $M = 0.63$  [0.62, 0.63]) rating AI as slightly less competent than men ( $M = 0.64$  [0.64, 0.64]),  $t(11398.71) = -7.94$ ,  $p < .001$ ,  $\Delta M = -0.02$  [-0.02, -0.01]. Women’s metaphors were also more anthropomorphic ( $M = 0.53$  [0.52, 0.53]) than men’s ( $M = 0.52$  [0.51, 0.52]),  $t(11136.83) = 3.06$ ,  $p = 0.007$ ,  $\Delta M = 0.01$  [0.00, 0.01], though this effect was small. Together, these results indicate that women and especially non-binary people conceptualize AI as less warm and competent, and in turn are less trusting of and less willing to adopt it.

Differences in rates of dominant metaphors give some insight into these patterns. Based on two-sample binomial tests, “robot” (men = 34.2%, women and non-binary = 65.8%),  $p < 0.001$ , and “genie” (men = 37.3%, women and non-binary = 62.7%),  $p < 0.001$ , were disproportionately more frequently provided by women and non-binary respondents. In contrast, men significantly more frequently provided highly competent or warm—



**Fig. 4 | Demographic differences in perceptions of AI (n = 11,790).** Each panel shows mean normalized perception scores (range [0, 1]) across demographic groups. Error bars denote 95% confidence intervals. Group comparisons were performed with two-sample t-tests between: men vs non-men (gender); white vs non-white (race/ethnicity); 18–37 vs 37+ years (age); individual contributors vs managers (work level); no college degree vs college degree or higher (education); using < 3 times versus > 3 times (frequency of use); and 0 vs > 0 AI tools heard of (number of tools heard of). Asterisks indicate significant differences (\* $p < 0.05$ , \*\* $p < 0.01$ ,

\*\*\* $p < 0.001$ ). Men report higher trust, adoption, warmth, and competence but lower anthropomorphism. Non-white participants exhibit higher trust, adoption, and anthropomorphism. Older participants trust and anthropomorphize AI more. Managers rate AI as warmer, more competent, and more trustworthy than individual contributors. Frequent users of AI show higher scores on all perception dimensions, whereas those unfamiliar with AI tools perceive AI as warmer but less competent.

but not necessarily anthropomorphic—metaphors such as “teacher” (men = 53.0%, women and non-binary = 47.0%),  $p = 0.009$ ; “search engine” (men = 50.5%, women and non-binary = 49.5%),  $p = 0.004$ ; library (men = 50.3%, women and non-binary = 49.7%),  $p = 0.035$ ; and “computer” (men = 49.4%, women and non-binary = 50.6%),  $p = .047$ , invoking entities that are familiar, helpful, and epistemically competent.

**Race/ethnicity**

Black participants had significantly higher trust ( $M = 0.66 [0.65, 0.67]$ ) and willingness to adopt AI ( $M = 0.58 [0.57, 0.59]$ ) than white participants

( $M = 0.56 [0.55, 0.56]$  for trust;  $M = 0.50 [0.50, 0.51]$  for adoption),  $t(2317.74) = -16.83, p < 0.001, \Delta M = -0.10 [-0.11, -0.09]$  for trust;  $t(2262.02) = -11.44, p < 0.001, \Delta M = -0.08 [-0.09, -0.06]$  for adoption. Asian ( $M = 0.58 [0.56, 0.59]$ ) and Hispanic participants ( $M = 0.59 [0.57, 0.60]$ ) also reported higher trust than white participants,  $t(1311.93) = -2.88, p = 0.049, \Delta M = -0.02 [-0.03, -0.01]$ ; and  $t(1016.54) = -3.81, p = 0.002, \Delta M = -0.03 [-0.05, -0.01]$ , respectively. In contrast, participants identifying as “Other” race/ethnicity ( $M = 0.52 [0.49, 0.54]$ ) had significantly lower trust,  $t(294.01) = 3.02, p = 0.035, \Delta M = 0.04 [0.01, 0.07]$ , suggesting that within the participants of color, perceptions vary widely.

Participants of color viewed AI as more warm, competent, and anthropomorphic than white participants. For warmth, Black ( $M = 0.55$  [0.54, 0.55]) and Asian ( $M = 0.55$  [0.54, 0.55]) participants rated AI higher than white ( $M = 0.53$  [0.53, 0.53]),  $t(2488.86) = -6.35$ ,  $p < 0.001$ ,  $\Delta M = -0.02$  [-0.02, -0.01], and  $t(1372.58) = -5.17$ ,  $p < 0.001$ ,  $\Delta M = -0.02$  [-0.03, -0.01], respectively. Competence followed a similar pattern: Black ( $M = 0.64$  [0.64, 0.65]) and Asian ( $M = 0.65$  [0.64, 0.65]) participants rated AI as slightly more competent than white ( $M = 0.63$  [0.62, 0.63]),  $t(2523.77) = -5.72$ ,  $p < 0.001$ ,  $\Delta M = -0.02$  [-0.02, -0.01]; and  $t(1418.09) = -6.14$ ,  $p < 0.001$ ,  $\Delta M = -0.02$  [-0.03, -0.01]. Anthropomorphism was also marginally higher among Black ( $M = 0.54$  [0.53, 0.54]) than white ( $M = 0.52$  [0.51, 0.52]),  $t(2504.47) = -5.56$ ,  $p < 0.001$ ,  $\Delta M = -0.02$  [-0.03, -0.01]. Two-sample binomial tests comparing each non-white racial group to white participants revealed significant differences in metaphor usage. “Search engine” metaphors were significantly more common among white participants (73.0% white, 95% CI [70.2%, 75.6%],  $p < 0.001$ ), while “genie” metaphors were more common among participants of color (58.2% white, 95% CI [53.9%, 62.5%],  $p = 0.009$ ).

### Age

Older people trust AI more and are more willing to adopt it, and view it as more anthropomorphic but less competent: Adults 38–53 reported higher trust than those aged 18–37 ( $M = 0.59$  [0.58, 0.60] vs 0.56 [0.56, 0.57]),  $t(7371.56) = -6.36$ ,  $p < 0.001$ ,  $\Delta M = -0.03$  [-0.04, -0.02]; and those 54+ were also higher than 18–37 ( $M = 0.59$  [0.58, 0.60]),  $t(1437.06) = -3.68$ ,  $p < 0.001$ ,  $\Delta M = -0.03$  [-0.04, -0.01]. For adoption, 38–53 exceeded 18–37 ( $M = 0.54$  [0.53, 0.54] vs 0.50 [0.50, 0.51]),  $t(7346.26) = -7.15$ ,  $p < 0.001$ ,  $\Delta M = -0.04$  [-0.04, -0.03], though the 54+ category did not reliably differ from 18–37 ( $t(1423.66) = -1.55$ ,  $p = 0.122$ ). Anthropomorphism was slightly higher among 54+ than 18–37 ( $M = 0.53$  [0.52, 0.54] vs 0.52 [0.52, 0.52]),  $t(1665.39) = -2.52$ ,  $p = 0.035$ ,  $\Delta M = -0.01$  [-0.02, -0.00]. For competence, 18–37 exceeded 54+ ( $M = 0.63$  [0.63, 0.64] vs 0.63 [0.62, 0.63]),  $t(1626.33) = 2.48$ ,  $p = 0.039$ ,  $\Delta M = 0.01$  [0.00, 0.02]. Using two-sample binomial tests, the dominant metaphor “friend” was significantly less common among 18–37 (44.8% [39.1%, 50.5%],  $p = 0.001$ ), “child” was more common among 18–37 (65.6% [61.0%, 70.0%],  $p = 0.004$ ).

### Work level

We find that individuals higher in the organizational hierarchy—specifically frontline managers and managers-of-managers—are significantly more likely to perceive AI as warm and competent, and to express greater trust in and willingness to adopt AI, compared to individual contributors. For trust, group means increased linearly from individual contributors ( $M = 0.548$  [0.542, 0.553]) to frontline managers ( $M = 0.616$  [0.608, 0.624]) to managers-of-managers ( $M = 0.654$  [0.641, 0.668]),  $t(5848.68) = -14.16$ ,  $p < .001$ ,  $\Delta M = -0.068$  [-0.078, -0.059] for the first contrast and  $t(1245.37) = -14.50$ ,  $p < .001$ ,  $\Delta M = -0.107$  [-0.121, -0.092] for the second. Similarly, for willingness to adopt, individual contributors ( $M = 0.490$  [0.484, 0.496]) scored lower than frontline managers ( $M = 0.561$  [0.553, 0.570];  $t(5799.46) = -13.89$ ,  $p < 0.001$ ) and managers-of-managers ( $M = 0.595$  [0.580, 0.610];  $t(1204.74) = -12.71$ ,  $p < 0.001$ ), again showing a clear linear trend. For warmth and competence, the pattern paralleled trust and adoption but with smaller effect sizes: warmth rose modestly from 0.530 to 0.542 (individuals → frontline managers;  $t(6697.57) = -5.06$ ,  $p < 0.001$ ,  $g = -0.11$ ), and competence from 0.627 to 0.639 ( $t(6960.69) = -5.39$ ,  $p < 0.001$ ). Taken together, these results indicate a graded relationship between organizational rank and positive AI perceptions: managers-of-managers are the most trusting and adoption-oriented, followed by frontline managers, with individual contributors expressing the least enthusiasm. Based on two-sample  $t$ -tests, individual contributors were more likely to describe AI using metaphors such as “brain”—term suggesting replacement of human cognitive labor (brain = 54.3% [51.3%, 57.3%],  $p = 0.024$ )—whereas managers tended to use “animal” (e.g., animal = 69.2% [62.4%, 75.3%],  $p = 0.084$ ), which confers less agency to AI.

### Education

We find no statistically significant differences across education levels. This may be because educational attainment reflects a broad range of knowledge and experiences not necessarily related to AI or technology experience, and because we do not differentiate between those currently in college versus those who did not enter, which may reflect different socioeconomic groups.

### Number of AI tools heard of

We find that participants who had not heard of any listed AI tools viewed AI as warmer and more anthropomorphic relative to those familiar with more tools. For warmth, those who had heard of 0 tools ( $M = 0.56$  [0.55, 0.56]) > 1–2 ( $M = 0.53$  [0.53, 0.53]),  $t(1915.59) = 7.60$ ,  $p < 0.001$ ,  $\Delta M = 0.02$  [0.02, 0.03]; and 0 > 3+ ( $M = 0.53$  [0.53, 0.53]),  $t(2562.11) = 7.33$ ,  $p < 0.001$ ,  $\Delta M = 0.03$  [0.02, 0.03]. For anthropomorphism, 0 > 3+,  $t(2395.55) = 2.40$ ,  $p = .049$ ,  $\Delta M = 0.01$  [0.00, 0.02]. This suggests that limited knowledge of AI may foster more relational or human-like conceptualizations, potentially due to unfamiliarity with its actual capabilities or limitations. In turn, this suggests that knowledge about, or lack thereof, shape participants’ metaphors and influences downstream attitudes regarding trust and adoption.

### Frequency of AI use

Participants who used AI more frequently had significantly higher trust, adoption, warmth, and competence scores than those who used AI rarely or not at all. For trust, frequent users (10+ times per week;  $M = 0.69$  [0.66, 0.73]) exceeded moderate users (3–10;  $M = 0.64$  [0.63, 0.64]),  $t(148.15) = -3.18$ ,  $p = 0.002$ ,  $\Delta M = -0.06$  [-0.09, -0.02], who in turn exceeded infrequent users (<3;  $M = 0.52$  [0.51, 0.52]),  $t(10\ 487.88) = -29.93$ ,  $p < 0.001$ ,  $\Delta M = -0.12$  [-0.13, -0.11]. For adoption, frequent users ( $M = 0.67$  [0.64, 0.71]) again exceeded both moderate ( $M = 0.58$  [0.57, 0.58]),  $t(148.90) = -5.24$ ,  $p < 0.001$ ,  $\Delta M = -0.10$  [-0.13, -0.06], and infrequent users ( $M = 0.45$  [0.45, 0.46]),  $t(10\ 478.90) = -28.73$ ,  $p < 0.001$ ,  $\Delta M = -0.12$  [-0.13, -0.12]. For warmth, frequent users ( $M = 0.56$  [0.54, 0.57]) were higher than moderate ( $M = 0.55$  [0.54, 0.55]),  $t(148.97) = -1.07$ ,  $p = 0.285$ , and both exceeded infrequent users ( $M = 0.52$  [0.52, 0.53]),  $t(10\ 787.56) = -10.08$ ,  $p < 0.001$ ,  $\Delta M = -0.02$  [-0.02, -0.02]. Competence followed a similar pattern: frequent users ( $M = 0.64$  [0.63, 0.66]) and moderate users ( $M = 0.64$  [0.64, 0.65]) both scored higher than infrequent users ( $M = 0.62$  [0.62, 0.62]); <3–3–10  $t(11\ 075.03) = -12.05$ ,  $p < 0.001$ ,  $\Delta M = -0.02$  [-0.03, -0.02]; <3–10+  $t(149.42) = -2.80$ ,  $p = 0.012$ ,  $\Delta M = -0.02$  [-0.04, -0.01]. This suggests that more frequent users not only trust and adopt AI more readily, but also conceptualize it in ways that are both human-like and agentic.

### Discussion

Identifying the dominant metaphors people use to understand AI reveals the public’s rich and complex understanding of these evolving, increasingly ubiquitous tools. Our work complements existing psychological approaches for understanding human-AI relationships through surveys and experiments (e.g., refs. 49,97,98, by leveraging large-scale analyses of metaphors through a scalable framework. Eliciting participants’ perceptions allows people of all levels of familiarity with AI to express how they feel about AI even if they do not have the technical language to articulate their thoughts<sup>99</sup>.

Using this approach, we observe evidence of a social shift where the US population increasingly ascribe anthropomorphic and warm qualities to AI. This movement from perceiving AI as a “tool” or a “calculator” towards seeing it as an “assistant” or “friend” aligns with recent marketing narratives from AI companies that increasingly frame their products as “co-pilots” or “companions”<sup>100</sup> – raising important questions about the benefits and harms of ascribing greater agency to these technologies<sup>83</sup>. Although our data contrast findings from several years ago indicating people’s resistance to viewing AI relationally as “friends”<sup>59</sup>, they align with more recent research indicating that people tend to view AI as more humanlike as they come to use it more<sup>101</sup>. It is also notable that fewer shifts were seen in perceptions of competence. An analysis of newspaper AI coverage from the last four decades found that popular narratives around AI tended to concern its perceived capabilities, such as the idea that it may develop into

“superintelligence”<sup>102</sup>. Together, our findings suggest that public perceptions of AI may be taking a relatively recent shift toward relational framing<sup>57,103</sup>.

Understanding the extent to which people believe that AI-based technologies have agency is increasingly important as anthropomorphism increases and the perceived gap between humans and AI continues to narrow<sup>56,104</sup>. Scholars have long known that people implicitly view technologies as having human-like qualities, a perception that can shift their behaviors and outcomes (e.g., the computers-as-social-actors framework<sup>105</sup>). When people perceive an entity – human or artificial – to be more similar to themselves, they tend to increase their mind perception<sup>106</sup>. Anthropomorphizing AI can make people ascribe more mental capacities, such as believing they can experience human feelings or drives (i.e., hunger, fear, pride) or have the agency to develop and act on its own judgment<sup>107,108</sup>. These perceptions can be strong predictors of individual interactions with AI. For example, people tend to like service robots that are perceived as having human feelings, and are also more likely to forgive them for mistakes<sup>109</sup>.

This social shift has important implications for ongoing efforts to appropriately calibrate trust in AI<sup>110</sup>. People tend to trust human-like AI systems more<sup>7,111</sup>, which raises both opportunities and risks. On the beneficial side, technological systems that have anthropomorphic qualities, such as the ability to have human-like conversations or human-like avatars, can be perceived as easier to use and more accessible to people of all levels of technical expertise<sup>112,113</sup>. However, this increase in trust may increase individuals’ willingness to disclose sensitive personal information<sup>114</sup> and their susceptibility to persuasion or deception<sup>57</sup>.

Crucially, our work highlights how individuals from different backgrounds perceive AI in substantively different ways. In line with past work<sup>20,26,36,115,116</sup>, we found that men tended to trust and hold more positive attitudes toward AI more than women. However, closer examination of the metaphors they used to describe AI revealed that women and non-binary people provided more anthropomorphic metaphors of AI – highlighting the importance of considering perceptions of AI through different approaches. Our findings also reveal unexpectedly higher levels of trust among older adults and US Americans of color. While previous research found that Hispanic and Asian individuals tend to report more positive experiences with AI while white and Black individuals viewed AI as more harmful<sup>26</sup>, our findings add nuance to these results by examining how diverse US Americans think and feel about AI in multifaceted ways. Older people may index more strongly on AI as potential companions, though further research is needed to understand this phenomenon. Our findings also align with evidence that lower-ranking employees are more likely to perceive AI as a threat to their job security or as a potential replacement. These findings point to the importance of identity in human-AI interactions and underscore the need to consider these differences in both design and policy.

### Limitations

This study has several limitations that should be considered when interpreting our results. First, we note that we chose to focus on examining the anthropomorphism, warmth, and competence of the metaphors we studied, given our research interests in examining perceptions of the similarity of AI to humans<sup>1,53</sup>. However, many other perceptual dimensions of metaphors of AI exist<sup>38</sup>. For example, people vary in their beliefs about the power and control they have over AI systems and may perceive AI as acting with differing degrees of agency<sup>104,117</sup>. These perceptions may also be reflected in the metaphors people use to describe AI, such as viewing it as a “tool” they are in control of using or a “thought partner” that is viewed as more like an intellectual equal<sup>17</sup>. While our research advances our understanding of how the US population’s mental models of AI have shifted in terms of their similarity to humans, future research should further investigate shifts in additional dimensions of AI perceptions to obtain a fuller picture of how people’s views of AI may be changing<sup>107</sup>.

Future research should examine how people perceive different kinds of AI-based technologies. This study focused on investigating broad public

conceptualizations of AI: when 90% of the population reports familiarity with AI, what exactly are they referring to? However, people may come to hold different beliefs about distinct tools, such as LLM-based chatbots like ChatGPT, virtual assistants like Alexa, and companions like Replika, that vary in their design and stated purpose<sup>111,118</sup>. Understanding shifts in public perception in specific AI tools is particularly important as these technologies continue to evolve and as more people gain familiarity with a greater variety of tools and usage increases<sup>119</sup>. In particular, individuals’ willingness to trust AI and adopt it in the future often depends on specific tools and the extent to which each aligns with their goals (see refs. 76,120).

Similarly, more research is needed to examine how individual differences in individuals’ knowledge of AI shapes the metaphors they use to conceptualize AI, and their trust and willingness to adopt it. While this study examined how participants’ awareness of AI tools and usage of these tools related to their perceptions of AI, it did not assess their AI literacy: their capacity to understand and apply AI tools to relevant tasks, to do so authentically by focusing on genuine communication, to take accountability for the reliability and validity of content generated with AI, and to maintain their agency by retaining decision-making rights and avoiding long-term dependency<sup>121</sup>. People with different levels of AI literacy should perceive AI differently, given previous work on algorithmic literacy indicating that having more in-depth knowledge and experience with complex socio-technical systems, like recommendation algorithms, changed how people viewed the role it played in their lives<sup>19</sup>, underscoring our findings that limited knowledge of AI may foster more relational or human-like conceptualizations, potentially due to unfamiliarity with its actual capabilities or limitations. Understanding how AI literacy relates to people’s metaphors of AI (i.e., “AI is a thief”, “AI is a copycat”) can help scholars identify communities that may be disproportionately, negatively affected by AI, even if they do not have the technical vocabulary to express these concerns fully.

Finally, we note that we prioritized collecting a large, representative dataset from each month in our sample. As a result, this may result in more variables being significant in our analyses. Effect sizes with betas and confidence intervals, such as those in Fig. 3, should be taken into consideration when interpreting these analyses.

### Conclusion

Our study examines the evolution of public perceptions of AI over time among the US American public through a novel dataset of 12,000 metaphors. Within the short timespan of just a year and a half, we see substantial shifts in how people think and feel about AI, highlighting the impact that the rapid adoption of systems like ChatGPT and Gemini have on our collective consciousness. Taken together, our findings reflect a societal shift in people seeing AI as being more human-like and warm, and increasingly distinct from other digital technologies. Our findings also provide evidence that people from different backgrounds conceptualize AI differently, which in turn influences downstream attitudes regarding trust and adoption. Increasing perceptions of AI as warm and human-like has important implications in the design, deployment, and regulation of AI systems. Designers may consider adopting interventions to reduce human-likeness in AI system outputs<sup>53</sup> and/or include disclaimers when concerns about over-reliance are salient<sup>122</sup>. In the domain of deployment, while perceptions of warmth and anthropomorphism may be beneficial in specific contexts—for example, therapeutic or assistive settings—we call for further research into when and where such design choices meaningfully enhance utility without exacerbating risks. For instance, important questions remain around acceptable practices in high-stakes applications, such as whether chatbots aimed at children may legally present themselves as “friends,” or what boundaries should exist in systems that function as social replacements<sup>21</sup>.

### Data availability

All data is available at <https://doi.org/10.17605/OSF.IO/P6KFC>. The dataset contains de-identified participant-level responses, including the metaphors, demographic variables, and derived scores and dominant metaphors used in

the analyses. The latter variables were computed by the authors based on the open-ended metaphor responses and subsequent coding and scoring procedures described in the Methods section. The dataset is publicly available and can be downloaded under a CC-BY 4.0 license.

### Code availability

All code is available at <https://doi.org/10.17605/OSF.IO/P6KFC>.

Received: 19 April 2025; Accepted: 26 November 2025;

Published online: 09 January 2026

### References

- Gilardi, F., Kasirzadeh, A., Bernstein, A., Staab, S. & Gohdes, A. We need to understand the effect of narratives about generative AI. *Nature Hum. Behav.*, 1–2. (2024).
- Brugman, B. C., Burgers, C. & Steen, G. J. Recategorizing political frames: a systematic review of metaphorical framing in experiments on political communication. *Ann. Int. Commun. Assoc.* **41**, 181–197 (2017).
- Jensen, T. Disentangling trust and anthropomorphism toward the design of human-centered AI systems. In International Conference on Human-Computer Interaction (pp. 41–58). Cham: Springer International Publishing. (2021).
- Kasirzadeh, A. & Gabriel, I. In conversation with artificial intelligence: aligning language models with human values. *Philos. Technol.* **36**, 27 (2023).
- Demir, K. & Güraksın, G. E. Determining middle school students' perceptions of the concept of artificial intelligence: A metaphor analysis. *Particip. Educ. Res.* **9**, 297–312 (2022).
- Roth, E. (2024). "ChatGPT now has over 300 million weekly users." <https://www.theverge.com/2024/12/4/24313097/chatgpt-300-million-weekly-users>.
- Glikson, E. & Woolley, A. W. Human trust in artificial intelligence: Review of empirical research. *Acad. Manag. Ann.* **14**, 627–660 (2020).
- Sartori, L. & Bocca, G. Minding the gap(s): public perceptions of AI and socio-technical imaginaries. *AI Soc.* **38**, 443–458 (2023).
- Hazarika, I. Artificial intelligence: opportunities and implications for the health workforce. *Int. Health* **12**, 241–245 (2020).
- Danaher, J. Toward an ethics of AI assistants: An initial framework. *Philos. Technol.* **31**, 629–653 (2018).
- Doshi, A. R. & Hauser, O. P. Generative AI enhances individual creativity but reduces the collective diversity of novel content. *Sci. Adv.* **10**, eadn5290 (2024).
- Kochhar, R. Which US workers are more exposed to AI on their jobs? Washington, DC, USA: Pew Research Center. (2023).
- Starke, C. et al. Risks and protective measures for synthetic relationships. *Nat. Hum. Behav.* **8**, 1834–1836 (2024).
- Kelly, S., Kaye, S. A. & Oviedo-Trespalacios, O. What factors contribute to the acceptance of artificial intelligence? A systematic review. *Telemat. Inform.* **77**, 101925 (2023).
- Wilczek, B., Thäsler-Kordonouri, S. & Eder, M. Government regulation or industry self-regulation of AI? Investigating the relationships between uncertainty avoidance, people's AI risk perceptions, and their regulatory preferences in Europe. *AI & Society*, 1–15. (2024).
- Jiang, J. A., Wade, K., Fiesler, C. & Brubaker, J. R. Supporting serendipity: Opportunities and challenges for human-AI collaboration in qualitative analysis. *Proc. ACM Hum.-Comput. Interact.* **5**, 1–23 (2021).
- Kim, T., Molina, M. D., Rheu, M., Zhan, E. S. & Peng, W. One AI does not fit all: A cluster analysis of the laypeople's perception of AI roles. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (pp. 1–20). (2023).
- Lee, A. Y., Katz, R. & Hancock, J. The role of subjective construals on reporting and reasoning about social media use. *Soc. Media + Soc.* **7**, 20563051211035350 (2021).
- DeVito, M. A. Adaptive folk theorization as a path to algorithmic literacy on changing platforms. *Proceedings of the ACM on Human-Computer Interaction*, **5** (CSCW2), 1–38. (2021).
- Rice, R. E. & Wu, M. Y. Difference in and influences on public opinion about artificial intelligence in 20 economies: Reducing uncertainty through awareness, knowledge, and trust. *Int. J. Commun.* **19**, 26 (2025).
- Dreksler, N. et al. What does the public think about AI? An overview of the public's attitudes towards AI and a resource for future research. (2025).
- Lakoff, G. & Johnson, M. The metaphorical structure of the human conceptual system. *Cogn. Sci.* **4**, 195–208 (1980).
- Wu, T. et al. A brief overview of ChatGPT: The history, status quo and potential future development. *IEEE/CAA J. Autom. Sin.* **10**, 1122–1136 (2023).
- Gentner, D. Structure-mapping: A theoretical framework for analogy. *Cogn. Sci.* **7**, 155–170 (1983).
- Thibodeau, P. H. & Boroditsky, L. Natural language metaphors covertly influence reasoning. *PLOS ONE* **8**, e52961 (2013).
- Paik, S., Novozhilova, E., Mays, K. K. & Katz, J. E. Who benefits from AI? Examining different demographics' fairness perceptions across personal, work, and public life. *Discov. Artif. Intell.* **5**, 39 (2025).
- Faverio, M. & Tyson, A. What the data says about Americans' views of artificial intelligence. (2023).
- Benko, A. & Lányi, C. S. History of artificial intelligence. In *Encyclopedia of Information Science and Technology, Second Edition* (pp. 1759–1762). IGI Global. (2009).
- Cheng, M., Gligorić, K., Piccardi, T. & Jurafsky, D. AnthroScore: A computational linguistic measure of anthropomorphism. In *Proceedings of the 18th Conference of the European Chapter of the Association for Computational Linguistics (Volume 1: Long Papers)* (pp. 807–825). (2024).
- Fraser, K. C., Nejadgholi, I. & Kiritchenko, S. Understanding and countering stereotypes: A computational approach to the stereotype content model. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)* (pp. 600–616). (2021).
- Fiske, S. T., Cuddy, A. J. & Glick, P. Universal dimensions of social cognition: Warmth and competence. *Trends Cogn. Sci.* **11**, 77–83 (2007).
- Mieczkowski, H., Liu, S. X., Hancock, J. & Reeves, B. Helping not hurting: Applying the stereotype content model and bias map to social robotics. In *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)* (pp. 222–229). IEEE. (2019).
- Jan, Z. et al. Artificial intelligence for industry 4.0: Systematic review of applications, challenges, and opportunities. *Expert Syst. Appl.* **216**, 119456 (2023).
- Ma, X., Hancock, J. T., Lim Mingjie, K. & Naaman, M. Self-disclosure and perceived trustworthiness of Airbnb host profiles. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing* (pp. 2397–2409). (2017).
- Troshani, I., Rao Hill, S., Sherman, C. & Arthur, D. Do we trust in AI? Role of anthropomorphism and intelligence. *J. Comput. Inf. Syst.* **61**, 481–491 (2021).
- Yang, S. et al. In AI we trust: The interplay of media use, political ideology, and trust in shaping emerging AI attitudes. *J. Mass Commun. Q.* 10776990231190868. (2023).
- Khadpe, P., Krishna, R., Fei-Fei, L., Hancock, J. T. & Bernstein, M. S. Conceptual metaphors impact perceptions of human-AI collaboration. *Proceedings of the ACM on Human-Computer Interaction*, **4**(CSCW2), 1–26. (2020).
- Li, Y. et al. Warmth, competence, and the determinants of trust in artificial intelligence: A cross-sectional survey from China. *International Journal of Human-Computer Interaction*, 1–15. (2024).

39. McKee, K. R., Bai, X. & Fiske, S. T. Humans perceive warmth and competence in artificial intelligence. *iScience*, **26**. (2023).
40. Zhou, K. et al. (2024). Rel-AI: An interaction-centered approach to measuring human–LM reliance. In *Proceedings of the 2025 Conference of the Nations of the Americas Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 1: Long Papers)* (pp. 11148–11167).
41. Lakoff, G. & Johnson, M. *Metaphors we live by*. University of Chicago Press. (1981).
42. Hendricks, R. K., Demjén, Z., Semino, E. & Boroditsky, L. Emotional implications of metaphor: Consequences of metaphor framing for mindset about cancer. *Metaphor Symb.* **33**, 267–279 (2018).
43. Tversky, A. & Kahneman, D. Judgment under uncertainty: Heuristics and biases: Biases in judgments reveal some heuristics of thinking under uncertainty. *Science* **185**, 1124–1131 (1974).
44. Thibodeau, P. H. Extended metaphors are the home runs of persuasion: Don't fumble the phrase. *Metaphor Symb.* **31**, 53–72 (2016).
45. Fluckiger, F. From world-wide Wide Web to Information Superhighway. *Comput. Netw. ISDN Syst.* **28**, 525–534 (1996).
46. Maglio, P. P. & Matlock, T. Metaphors we surf the web by. In *Workshop on Personalized and Social Navigation in Information Space* (pp. 1–9). (1998).
47. Ytre-Arne, B. & Moe, H. Folk theories of algorithms: Understanding digital irritation. *Media Cult. Soc.* **43**, 807–824. (2021).
48. Brewer, P. R., Bingaman, J., Paintsil, A., Wilson, D. C. & Dawson, W. Media use, interpersonal communication, and attitudes toward artificial intelligence. *Sci. Commun.* **44**, 559–592 (2022).
49. Liehner, G. L., Biermann, H., Hick, A., Brauner, P. & Ziefle, M. Perceptions, attitudes and trust towards artificial intelligence—An assessment of the public opinion. *Artif. Intell. Soc. Comput.* **72**, 32–41 (2023).
50. Blut, M., Wang, C., Wunderlich, N. V. & Brock, C. Understanding anthropomorphism in service provision: A meta-analysis of physical robots, chatbots, and other AI. *J. Acad. Mark. Sci.* **49**, 632–658 (2021).
51. Kaplan, A. D., Kessler, T. T., Brill, J. C. & Hancock, P. A. Trust in artificial intelligence: Meta-analytic findings. *Hum. Factors* **65**, 337–359 (2023).
52. Epley, N., Waytz, A. & Cacioppo, J. T. On seeing human: A three-factor theory of anthropomorphism. *Psychol. Rev.* **114**, 864 (2007).
53. Cheng, M., Blodgett, S. L., DeVrio, A., Egede, L. & Olteanu, A. Dehumanizing machines: Mitigating anthropomorphic behaviors in text generation systems. In *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)* (pp. 25923–25948). (2025).
54. DeVrio, A., Cheng, M., Egede, L., Olteanu, A. & Blodgett, S. L. A taxonomy of linguistic expressions that contribute to anthropomorphism of language technologies. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems* (pp. 1–18). (2025).
55. Reeves, B. & Nass, C. *The media equation: How people treat computers, television, and new media like real people*. Cambridge University Press. (1996).
56. Guzman, A. L. Ontological boundaries between humans and computers and the implications for human–machine communication. *Hum.-Mach. Commun.* **1**, 37–54 (2020).
57. Peter, S., Riemer, K. & West, J. D. The benefits and dangers of anthropomorphic conversational agents. *Proc. Natl. Acad. Sci.* **122**, e2415898122 (2025).
58. Cheng, M., DeVrio, A., Egede, L., Blodgett, S. L. & Olteanu, A. “I am the one and only, your cyber BFF”: Understanding the impact of GenAI requires understanding the impact of anthropomorphic AI. In *ICLR Blogposts 2025*. (2025).
59. Philipsen, R., Brauner, P. M., Biermann, H. & Ziefle, M. C. I am what I am—Roles for artificial intelligence from the users' perspective. *Universitätsbibliothek der RWTH Aachen*. (2022).
60. Cuddy et al. Stereotype content model across cultures: Towards universal similarities and some differences. *Br. J. Soc. Psychol.* **48**, 1–33 (2009).
61. Afroogh, S., Akbari, A., Malone, E., Kargar, M. & Alambeigi, H. Trust in AI: Progress, challenges, and future directions. *Hum. Soc. Sci. Commun.* **11**, 1–30 (2024).
62. Choudhury, A. & Shamszare, H. Investigating the impact of user trust on the adoption and use of ChatGPT: Survey analysis. *J. Med. Internet Res.* **25**, e47184 (2023).
63. Abercrombie, G., Curry, A. C., Dinkar, T., Rieser, V. & Talat, Z. Mirages: On anthropomorphism in dialogue systems. In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing* (pp. 4776–4790). (2023).
64. Chiesurin, S. et al. The dangers of trusting stochastic parrots: Faithfulness and trust in open-domain conversational question answering. *arXiv Prepr. arXiv* **2305**, 16519 (2023).
65. Birhane, A. Algorithmic injustice: A relational ethics approach. *Patterns*, **2**. (2021).
66. Bianchi, F. et al. (2023). Easily accessible text-to-image generation amplifies demographic stereotypes at large scale. In *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency* (pp. 1493–1504).
67. Kalluri, P. R. et al. Computer-vision research powers surveillance technology. *Nature*, 1–7. (2025).
68. Chien, J. & Danks, D. Beyond behaviorist representational harms: A plan for measurement and mitigation. In *Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency* (pp. 933–946). (2024).
69. Lukyanenko, R., Maass, W. & Storey, V. C. Trust in artificial intelligence: From a Foundational Trust Framework to emerging research opportunities. *Electron. Mark.* **32**, 1993–2020 (2022).
70. Baker, R. et al. Summary report of the AAPOR task force on non-probability sampling. *J. Surv. Stat. Methodol.* **1**, 90–143 (2013).
71. Pasek, J. When will nonprobability surveys mirror probability surveys? Considering types of inference and weighting strategies as criteria for correspondence. *Int. J. Public Opin. Res.* **28**, 269–291 (2016).
72. Mercer, A., Lau, A. & Kennedy, C. For weighting online opt-in samples, what matters most? Pew Research Center. (2018).
73. Gramlich, J. What the 2020 electorate looks like by party, race and ethnicity, age, education and religion. Pew Research Center. <https://www.pewresearch.org/short-reads/2020/10/26/what-the-2020-electorate-looks-like-by-party-race-and-ethnicity-age-education-and-religion/>.
74. Jakesch, M., French, M., Ma, X., Hancock, J. T. & Naaman, M. AI-mediated communication: How the perception that profile text was written by AI affects trustworthiness. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (pp. 1–13). (2019).
75. Mayer, R. C. An Integrative Model of Organizational Trust. *Academy of Management Review*. (1995).
76. Gillespie, N., Lockey, S., Ward, T., Macdade, A. & Hassed, G. Trust, attitudes and use of artificial intelligence. (2025).
77. Goodrich, B., Fenton, M., Penn, J., Bovay, J. & Mountain, T. Battling bots: Experiences and strategies to mitigate fraudulent responses in online surveys. *Appl. Econ. Perspect. Policy* **45**, 762–784 (2023).
78. Zhang, S., Xu, J. & Alvero, A. J. (2025). Generative ai meets open-ended survey responses: Research participant use of AI and homogenization. *Sociol. Methods Res.* 00491241251327130.
79. Douglas, B. D., Ewell, P. J. & Brauer, M. Data quality in online human-subjects research: Comparisons between MTurk, Prolific,

- CloudResearch, Qualtrics, and SONA. *PLoS one* **18**, e0279720 (2023).
80. Barends, A. J. & De Vries, R. E. Noncompliant responding: Comparing exclusion criteria in MTurk personality research to improve data quality. *Personal. Individ. Differ.* **143**, 84–89 (2019).
  81. Nardon, L. & Hari, A. Sensemaking through metaphors: The role of imaginative metaphor elicitation in constructing new understandings. *Int. J. Qualit. Methods*, **20**, 16094069211019589. (2021).
  82. Reimers, N. & Gurevych, I. Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)* (p. 3982). Association for Computational Linguistics. (2019).
  83. Kirk, H. R. et al. The PRISM alignment dataset: What participatory, representative and individualised human feedback reveals about the subjective and multicultural alignment of large language models. *Adv. Neural Inf. Process. Syst.* **37**, 105236–105344 (2024).
  84. McInnes, L., Healy, J. & Melville, J. UMAP: Uniform manifold approximation and projection for dimension reduction. *arXiv Prepr. arXiv* **1802**, 03426 (2018).
  85. McInnes, L., Healy, J. & Astels, S. hdbSCAN: Hierarchical density based clustering. *J. Open Source Softw.* **2**, 205 (2017).
  86. Grootendorst, M. BERTopic: Neural topic modeling with a class-based TF-IDF procedure. *arXiv Prepr. arXiv* **2203**, 05794 (2022).
  87. Charmaz, K. Grounded theory. *Qualit. Psychol.: A Pract. Guide Res. methods* **3**, 53–84 (2015).
  88. Corbin, J. & Strauss, A. (2014). *Basics of qualitative research: Techniques and procedures for developing grounded theory*. Sage Publications.
  89. Hill, C. E. et al. Consensual qualitative research: an update. *J. Counsel. Psychol.* **52**, 196 (2005).
  90. Lucy, L., Tadimet, D. & Bamman, D. Discovering Differences in the Representation of People using Contextualized Semantic Axes. In *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing*. (2022).
  91. Cheng, M., Durmus, E. & Jurafsky, D. Marked personas: using natural language prompts to measure stereotypes in language models. In *The 61st Annual Meeting Of The Association For Computational Linguistics*. (2023).
  92. Crum, A. J., Salovey, P. & Achor, S. Rethinking stress: the role of mindsets in determining the stress response. *J. Personal. Soc. Psychol.* **104**, 716 (2013).
  93. Lee, A. Y. & Hancock, J. T. Social media mindsets: A new approach to understanding social media use and psychological well-being. *J. Comput.-Mediat. Commun.* **29**, zmad048 (2024).
  94. Cohn, M. et al. Believing anthropomorphism: examining the role of anthropomorphic cues on trust in large language models. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems* (pp. 1–15). (2024).
  95. Samuelson, P. Generative AI meets copyright. *Science* **381**, 158–161 (2023).
  96. Goetze, T. S. AI art is theft: Labour, extraction, and exploitation: Or, on the dangers of stochastic Pollocks. In *Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency* (pp. 186–196). (2024).
  97. Rosen, L. D., Carrier, L. M. & Cheever, N. A. Facebook and texting made me do it: Media-induced task-switching while studying. *Comput. Hum. Behav.* **29**, 948–958 (2013).
  98. Araujo, T., Helberger, N., Kruike-meier, S. & De Vreese, C. H. In AI we trust? Perceptions about automated decision-making by artificial intelligence. *AI Soc.* **35**, 611–623 (2020).
  99. Seering, J., Kaufman, G. & Chancellor, S. Metaphors in moderation. *N. Media Soc.* **24**, 621–640 (2022).
  100. Biermann, O. C., Ma, N. F. & Yoon, D. From tool to companion: Storywriters want AI writers to respect their personal values and writing strategies. In *Proceedings of the 2022 ACM Designing Interactive Systems Conference*, (pp. 1209–1227). (2022).
  101. Cardon, P. W. & Marshall, B. Can AI be your teammate or friend? Frequent AI users are more likely to grant humanlike roles to AI. *Bus. Profess. Commun. Q.* **87**, 654–669 (2024).
  102. Sun, S., Zhai, Y., Shen, B. & Chen, Y. Newspaper coverage of artificial intelligence: A perspective of emerging technologies. *Telemat. Inform.* **53**, 101433 (2020).
  103. Chaturvedi, R., Verma, S., Das, R. & Dwivedi, Y. K. Social companionship with artificial intelligence: Recent trends and future avenues. *Technol. Forecast. Soc. Change* **193**, 122634 (2023).
  104. Sundar, S. S. Rise of machine agency: A framework for studying the psychology of human–AI interaction (HAI). *J. Comput.-Mediat. Commun.* **25**, 74–88 (2020).
  105. Nass, C., Steuer, J. & Tauber, E. R. Computers are social actors. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems* (pp. 72–78). (1994).
  106. Waytz, A., Gray, K., Epley, N. & Wegner, D. M. Causes and consequences of mind perception. *Trends Cogn. Sci.* **14**, 383–388 (2010).
  107. Gray, H. M., Gray, K. & Wegner, D. M. Dimensions of mind perception. *Science* **315**, 619–619 (2007).
  108. Malle, B. F., Bello, P. & Scheutz, M. Requirements for an artificial agent with norm competence. In *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society* (pp. 21–27). (2019).
  109. Yam, K. C. et al. Robots at work: People prefer—and forgive—service robots with perceived feelings. *J. Appl. Psychol.* **106**, 1557 (2021).
  110. Schlicker, N. et al. How do we assess the trustworthiness of AI? Introducing the trustworthiness assessment model (TrAM). *Comput. Hum. Behav.* **170**, 108671 (2025).
  111. Hoff, K. A. & Bashir, M. Trust in automation: Integrating empirical evidence on factors that influence trust. *Hum. factors* **57**, 407–434 (2015).
  112. Wang, D. & Zhang, S. Large language models in medical and healthcare fields: applications, advances, and challenges. *Artif. Intell. Rev.* **57**, 299 (2024).
  113. Lim, J. H., Kwon, S., Yao, Z., Lalor, J. P. & Yu, H. *Large Lang. Model-based Role-Play. Pers. Med. Jarg. Extract. arXiv Prepr. arXiv* **2408**, 05555 (2024).
  114. Zhang, S. et al. “Ghost of the past”: identifying and resolving privacy leakage from LLM’s memory through proactive user interaction. *arXiv Prepr. arXiv* **2410**, 14931 (2024).
  115. Zhang, B. & Dafoe, A. Artificial intelligence: American attitudes and trends. Available at SSRN 3312874. (2019).
  116. McClain, C. Americans’ use of ChatGPT is ticking up, but few trust its election information. (2024).
  117. Lee, A. Y., Ellison, N. B. & Hancock, J. T. To use or be used? The role of agency in social media use and well-being. *Front. Comput. Sci.* **5**, 1123323 (2023).
  118. Bedué, P. & Fritzsche, A. Can we trust AI? An empirical investigation of trust requirements and guide to successful AI adoption. *J. Enterp. Inf. Manag.* **35**, 530–549 (2022).
  119. Maslej, N. et al. Artificial Intelligence Index Report 2025. *arXiv Prepr. arXiv* **2504**, 07139 (2025).
  120. Novozhilova, E., Mays, K., Paik, S. & Katz, J. E. More capable, less benevolent: Trust perceptions of AI systems across societal contexts. *Mach. Learn. Knowl. Extract.* **6**, 342–366 (2024).
  121. Cardon, P., Fleischmann, C., Aritz, J., Logemann, M. & Heidewald, J. The challenges and opportunities of AI-assisted writing: Developing AI literacy for the AI age. *Bus. Profess. Commun. Q.* **86**, 257–295 (2023).

122. Bo, J. Y., Kumar, H., Liut, M. & Anderson, A. Disclosures & disclaimers: Investigating the impact of transparency disclosures and reliability disclaimers on learner-LLM interactions. In *Proceedings of the AAAI Conference on Human Computation and Crowdsourcing* (Vol. 12, pp. 23–32). (2024).

### Acknowledgements

We extend thanks to all of the participants who shared their data with us to make the study possible and the reviewers whose feedback strengthened our work. In addition, we would like to thank the members of the Social Media Lab, including Sunny Liu, Harry Yan, Ryan C. Moore, and Ronald Robertson, for their feedback on this paper. AYL is supported by the Stanford Mark & Mary Stevens Interdisciplinary Graduate Fellowship. Data collection was funded by BetterUp. The funder provided support for the research and the conceptualization of the data collection, but did not have input into data analyses, reporting, or manuscript-writing.

### Author contributions

A.Y.L., J.H., A.L., K.R. and K.N. designed the study, selected measures, and collected the data. M.C. designed and implemented all computational analyses. M.C. created visualizations. M.C. A.Y.L., J.H. wrote the paper. J.H. and K.N. supervised the project.

### Competing interests

Authors with BetterUp affiliations are paid employees of BetterUp Inc.; A. Lee and J. Hancock are paid consultants within BetterUp Labs at BetterUp, Inc. MC does not have any affiliation with BetterUp and has no conflicts of interests to declare.

### Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s44271-025-00376-6>.

**Correspondence** and requests for materials should be addressed to Myra Cheng or Angela Y. Lee.

**Peer review information** *Communications Psychology* thanks the anonymous reviewers for their contribution to the peer review of this work. A peer review file is available. Primary Handling Editors: Yafeng Pan and Jennifer Bellingtier.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2026