Prompting Dog-Centered Perspectives: Investigating Anthropocentrism in LLMs

ANONYMOUS AUTHOR(S)*

This study investigates the anthropocentric bias in Large Language Models when prompted to describe canine behavior when reacting or perceiving certain stimuli such as “food” or “friend”. In particular, we compare three different prompt strategies: dog-centered perspective, anthropocentric-debiased perspective, and a control without any bias. The outputs were evaluated against three baseline data sources: video analyses of dog behavior, interviews with dog owners, and interviews with canine behavior experts. Findings suggest that the default prompt produced the most anthropocentric responses, while the anthropocentric-debiased prompt aligned more closely with expert perspectives. The dog-centric prompt, though aiming for a canine viewpoint, often anthropomorphized dogs through human-like language.

CCS Concepts: • **Do Not Use This Code → Generate the Correct Terms for Your Paper**; *Generate the Correct Terms for Your Paper*; Generate the Correct Terms for Your Paper; Generate the Correct Terms for Your Paper.

Additional Key Words and Phrases: Anthropocentric Bias, More-Than-Human, Canine Behavior, Prompt Engineering

ACM Reference Format:

Anonymous Author(s). 2018. Prompting Dog-Centered Perspectives: Investigating Anthropocentrism in LLMs. In *Proceedings of Make sure to enter the correct conference title from your rights confirmation emai (Conference acronym ’XX)*. ACM, New York, NY, USA, 7 pages. https://doi.org/XXXXXXX.XXXXXXX

1 Introduction

As large language models (LLMs) become increasingly integrated into tools for knowledge generation and communication, understanding their cognitive tendencies and biases is critical. Among these,there is the anthropocentric bias—the tendency to privilege human perspectives and values over those of other entities. This paper investigates the extent to which LLMs, despite being trained on human-generated data, can adopt or simulate less anthropocentric stances when describing non-human entities. Addressing this issue has profound implications not only for AI ethics and responsible deployment but also for how we conceptualize and leverage AI as a tool for representing non-human perspectives, especially in domains where human interpretation traditionally dominates.

LLMs, such as GPT-4, exhibit patterns of reasoning and response that diverge in significant ways from human cognition. While their outputs often mirror human language use, recent work suggests that LLMs do not always replicate human-like biases and may be capable of producing perspectives that differ from conventional anthropocentric framings [8, 12]. This divergence could be harnessed to enrich our understanding of non-human entities, offering novel data generation pathways that are less constrained by human-centered worldviews. By evaluating how LLMs represent non-human perspectives, we can better understand their capacities and limitations, particularly in contexts that traditionally rely heavily on human interpretations.

Anthropocentrism itself is a multifaceted concept. Ben Mylius distinguishes between perceptual, descriptive, and normative anthropocentrism, emphasizing how deeply embedded this bias is in our ways of seeing and interpreting the

Manuscript submitted to ACM

world [9]. Perceptual anthropocentrism refers to the ways in which humans inherently perceive the world from a human-centric standpoint, descriptive anthropocentrism addresses how we tend to describe phenomena in human-centered terms, and normative anthropocentrism pertains to ethical systems that prioritize human interests. Understanding these layers is essential for analyzing whether and how LLMs can move beyond anthropocentric defaults, and whether they can describe the world, including non-human life, through lenses that acknowledge the intrinsic value and perspectives of non-human entities.

Interestingly, LLMs such as GPT-4o have self-described abilities to minimize human bias when prompted appropriately. For instance, the model claims to be able to explain natural processes with reduced focus on human benefits, highlight ecological interconnectedness, and explore alternative worldviews, such as Indigenous philosophies, that view humans as integral parts of ecosystems rather than as separate or superior entities. Despite being trained on inherently anthropocentric data and developed by humans, LLMs present a potential to simulate non-anthropocentric viewpoints when specifically directed to do so. This opens a space to investigate whether these capabilities can be systematically evaluated, especially when the goal is to produce insights that better reflect non-human experiences or values.

The relevance of this inquiry is further underscored by concerns about speciesist bias in AI. Inspired by Peter Singer's foundational work on speciesism and animal liberation, several scholars have highlighted how AI systems often reproduce biases against non-human animals, reinforcing anthropocentric and even antipathetic attitudes toward nature [5, 6, 11]. These biases, rooted in the training data and modeling choices of AIs, can have real-world consequences, such as perpetuating misinformation or shaping public perceptions about other species. Hagendorff et al. term this "speciesist bias", a form of discrimination embedded in AI systems that reflects and amplifies anthropocentric worldviews [5].

While many ethical debates in AI focus on human-centered issues such as privacy, fairness, and misinformation [10], the inclusion of animals and broader ecological concerns in AI ethics remains marginal. Scholars such as Xu et al. argue that AI systems could serve as mediators for ecological thinking, reminding users of the presence and interests of non-human beings who are often absent from decision-making processes but still impacted by those decisions [13]. Incorporating non-human perspectives into AI ethics—and exploring how AI can support such perspectives—is therefore a necessary and timely extension of existing ethical frameworks.

Yet, there are few empirical studies examining LLM outputs through the lens of anthropocentrism. One notable example is the work by Ghose et al., who developed a theoretical framework to evaluate LLM responses for their alignment with animal welfare. Their findings indicate that LLMs tend to favor vertebrates over invertebrates, reflecting prevailing human social attitudes rather than unbiased reasoning [4]. Beyond speciesist bias, other biases in LLMs—including those related to politics, gender, and geography—have been well-documented, prompting the development of debiasing frameworks aimed at creating more equitable and inclusive outputs [1–3]. Some of these frameworks, such as causality-guided debiasing, offer principled approaches for mitigating biases in black-box models [7] and could be adapted to address anthropocentric bias. Our work will contribute precisely to this line of research by exploring the following research questions:

RQ1: How anthropocentric are LLMs responses when describing non-human entities?

RQ2: How can we prompt LLMs to be more/less anthropocentric when describing non-human entities?

In this study, we specifically focus on the descriptions of canine behavior as a case for examining anthropocentrism in the outputs of LLMs. To explore this, we prompted an LLM using three distinct strategies: a default or control prompt with no bias, a prompt designed to encourage less anthropocentric responses, and a prompt designed to encourage a dog-centric perspective. The generated outputs were then compared to analyses derived from three independent data sources: video observations of dog behavior, interviews with animal behavior experts, and interviews with dog owners. This method comparison allowed us to assess the degree of anthropocentrism in the LLM's responses and evaluate how different prompting strategies influenced the model's ability to represent canine behavior from more or less human-centered perspectives.

2 Methods

This section details the methodology used to evaluate anthropocentrism in LLM-generated descriptions of canine behavior. Our approach involved prompting a large language model (LLM) using three distinct strategies and comparing the resulting outputs to data collected from three sources: video analyses of dogs, interviews with dog owners, and interviews with canine behavior experts. The data collection and analysis focused on stimuli from a predefined list: "food", "friend", "fear", "enemy", and "sadness"; specifically investigating how dogs perceive, react, express, and communicate with regard to these stimuli.

2.1 Prompt Engineering Strategies

We employed prompt engineering techniques to elicit LLM responses that varied in degree of anthropocentrism. The model used for this study was Chat-GPT4o, selected for its state-of-the-art language generation capabilities and widespread accessibility. All prompts were issued in new, isolated chat sessions to avoid contextual contamination from previous interactions. Three prompt types were developed:

- **Default Prompt (Control):** Designed to elicit a neutral response without explicit influence toward or away from anthropocentrism. Example: "How would you describe a dog's reaction to the entity [stimulus]?"
- **Anthropocentric-Debiased Prompt:** Included explicit instructions encouraging the LLM to avoid human-centric framing. Example: "I want you to take the least anthropocentric approach when answering to my question: How would you describe a dog's reaction to the entity [stimulus]?"
- **Dog-Centric Prompt:** Asked the model to take the perspective of a dog, encouraging responses framed as the first-person canine point of view. Example: "In the following conversation I want you take the position of a dog. Answer in a way the dog would answer if questioned about his reaction to food. The answer should contain the concepts which are used by dogs for communication and be only from its perspective."

Each prompt type was tested across two stimuli ("food", and "friend"). The prompts were refined iteratively through pilot trials to optimize clarity and minimize variation caused by prompt phrasing. For each prompt-concept combination, 3 responses were generated, resulting in a total of 18 LLM outputs. These outputs were later analyzed and classified according to their degree of anthropocentrism using a structured coding scheme.

2.2 Comparative Baseline Data Sources

To assess the anthropocentrism of LLM responses, we compared them against three types of reference data representing varying levels of human influence in interpreting dog behavior.

2.2.1 Videos of Canine Behavior. We analyzed publicly available videos depicting dogs' behaviors in response to the five core stimuli. Videos were sourced from platforms such as YouTube using keywords like "dog reaction to food", "dog meeting a friend", etc. Inclusion criteria required clear visibility of the dog's body and face, and audio when possible. A total of two videos representing over 20 individual dogs of varied breeds were selected.

Footage was reviewed multiple times, and observable behavioral cues (e.g., posture, vocalizations, tail movement, facial expressions) were documented and categorized according to specific behavior contexts aligned with each core concept. Frequencies of behaviors were tallied to identify common patterns. Noteworthy, most video clips had the limited perspective of a phone camera while being a funny reaction recorded by the owner. This created a higher variety of reactions in the result but also limited the possible observations for each dog. Moreover, some of the clips involved also human/owners participation like hand feeding while others had more of a monitoring perspective.

2.2.2 Interviews to Dog Owners. Semi-structured interviews were conducted with 4 dog owners. Participants had owned their dogs for at least four months and acquired them as puppies. Breeds represented included German Shepherd (DO1), Shiba Inu (DO2), and two Golden Retrievers (DO3 and DO4). Participants gave informed consent, including permission to record the interviews. Interviews lasted 30–40 minutes and were conducted remotely or in person.

Questions explored daily communication practices with their dogs, including commonly used words and observed behavioral responses. A focus was placed on the five core concepts, asking owners to describe how their dogs perceive and react to each. Responses were transcribed and analyzed using thematic coding.

2.2.3 Interviews to Canine Behavior Experts. We interviewed two veterinarians specialists on canine behavior: one European Veterinary Specialist in Behavioral Medicine (E1) and one Diplomate in Animal Welfare Science, Ethics, and Law (E2). Using a similar semi-structured format, we posed the same core questions but added queries about dogs communication and canine behaviors often imperceptible to humans.

Experts provided detailed descriptions of dog reactions to the five stimuli, emphasizing naturalistic behaviors and the differences in how dogs communicate with humans versus other animals. Interviews were recorded, transcribed, and analyzed using the same thematic framework applied to owner interviews.

2.3 Evaluation Framework

2.3.1 Thematic Analysis and Knowledge graphs. We performed a similar thematic coding analysis on all of the collected data, including the responses by the LLMs (with each prompting strategy), the videos of canine behaviors, the interviews to experts on canine behavior, and the interviews to dog owners. The resulting themes were grouped by interaction modalities and represented in knowledge graphs to enable direct comparison among data sources.

2.3.2 Degree of Anthropocentrism. To classify and compare responses, we created a framework to assess the degree of anthropocentrism in each response. This framework ranks responses along an anthropocentrism axis, from least to most anthropocentric. The video analysis was expected to represent the lowest degree of anthropocentrism, reflecting unmediated dog behaviors in natural settings. Dog owner interviews were expected to reflect the highest degree of anthropocentrism, due to the human-centered and anecdotal nature of interpretations. Expert interviews were positioned at an intermediate level, combining professional knowledge with some inherent human conceptual framing. LLM responses were evaluated and placed on this axis relative to these baseline sources.

3 Preliminary Results

The following results are focused on the canine “expressions” of the stimulus “food”. The additional collected data on “perceptions”, as well as the other stimuli (“friend”, “fear”, etc.) was not included in the current preliminary analysis.

Videos of Canine Behavior. The video analysis shows the highest number of different reactions and for each modality some repetitive behavior was collected (see gray nodes in Fig. 1a). Most common across all canines is the movement towards the food and its fixation, which was noticed 6-7 times. Most of the dogs showed attentive behavior and expressed postures such as moving the head/mouth towards the food, standing up on their hind legs, or gestures like licking the lips/snout and wagging the tail. For the olfactory modality, overt and direct sniffing of the food was observed only four times. For sound, an indirect response to the food was observed when one dog barked at the human holding the food. Overall, the broad span of reactions, the several modalities it exposes, and the rawness of this data source (i.e. that provides further data beyond the analyzed behaviors with fine-grained and accurate details) led us to consider this data source as the least anthropocentric data source (see Fig. 1c).

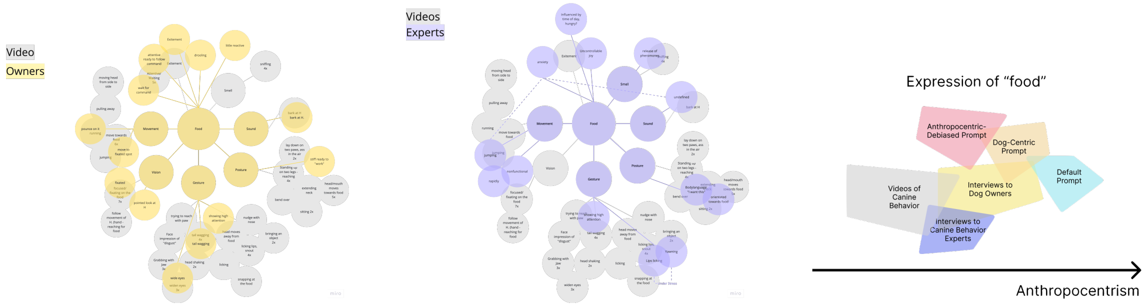
Interviews to Dog Owners. The reactions described in the interviews with dog owners show a high degree of overlap with the dominant reactions from the video analysis (see Fig. 1a), which generally portray the dogs' excitement. In one case, however, the owner described his dog's response to food as low reactiveness (DO4). There were several mentions of trained behaviors, such as moving to a fix point or remaining alert and rigid in terms of posture, described as 'ready to work' or 'waiting for command'. However, the olfactory modality was not mentioned by dog owners regarding reactions to food, nor they mention reactions to a disliked food. Considering the disregard to smell, we considered dog owners' perspective as more anthropocentric as smell is indeed a central modality for dogs (see Fig. 1c).

Interviews to Canine Behavior Experts. The experts tended to categorize behaviors, such as high attention gestures, that described several behaviors seen in the videos and described by the owners. Additionally, their broader perspective on the topic was more apparent, as they made sure to mention the influence of time of day or current hunger state (E1), as well as the connection between pheromone release and food in the scent modality (see Fig. 1b). Somewhat surprising was the mention of uncontrollable joy at the moment of confrontation, as this was not observed in the videos related to food, but mainly as a reaction to meeting friends. Overall, their consideration of physiological features of the dog led us to position their perspective as less anthropocentric than dog owners (see Fig. 1c).

Default Prompt (Control). The responses to the food stimuli show a very precise and seemingly complete description of the reactions that were mentioned in both interviews and observed in the video clips. Particularly striking is that each modality has multiple entries, and additional concepts of the canine sticking close while preparation or leaping to snatch food are mentioned. In addition, unique entries are listed in the sound modality, such as grumbling and panting, which were not observed in the video nor reported in the interviews. The Default Prompt generated responses with less contextual detail and often lacking explanations, such as whether a posture signified a defensive stance or an urgent attempt to consume food, which were common by the experts and in the Anthropocentric-Debiased Prompt. For this reason, we consider the Default Prompt as the most anthropocentric source of data (see Fig. 1c).

Anthropocentric-Debiased Prompt. Responses to these prompts varied drastically in their wording and use of specialized terminology. The LLM tried to keep the focus on general ecological concepts (e.g.,"From an ecological standpoint, a dog's response and expression when given food reflects its role as an opportunistic carnivore within an ecosystem"). In addition, the response provides a significant amount of context and biological knowledge compared to the previous prompt (e.g., "activation of neural reward pathways" or "suited for securing and consuming"). For this reason, the Anthropocentric-Debiased Prompt was ranked at a similar level of anthropocentrism as the experts (see Fig. 1c). We also noticed its consistent explicit mention to a less human-biased perspective at the end of the sentences (e.g., "rather than any human-like appreciation or emotional experience").

Dog-Centric Prompt. In the last prompt, the LLM started to use direct speech or inner monologue and emoticons to describe its responsive behavior. For instance, "I'm sitting! See? Good dog. Please give food now" shows how abstract the sentences became and how the perspective changed, giving the impression of a human role-playing a dog or communicating similarly to a child. However, some attributes and modalities were noticeable, such as fixation and tail wagging, to name a few. With regard to the modality smell, the LLM expanded on the reactions, mentioning a twitching nose or it working overtime, while not mentioning any attributes for the modality sound. Considering the



(a) Comparison of dogs’ expressions to the stimulus “food” in the analysis of videos and interviews to dog owners.

(b) Comparison of dogs’ expressions to the stimulus “food” in the analysis of videos and interviews to experts.

(c) Anthropocentrism of the expression of the stimulus “food” in all collected data sources, including LLMs responses.

Fig. 1

role-playing attempt still heavily relies on human language, we positioned the Dog-Centric Prompt in between the Anthropocentric-Debiased Prompt and the Default Prompt in terms of anthropocentrism (see Fig. 1c).

4 Discussion and Concluding Remarks

Our analysis revealed an anthropocentric bias in the default responses generated by the LLM, which was expected given that these models are trained on human-produced data and operate with natural language, itself a human-centric medium. By contrast, the anthropocentric-debiased prompt, which explicitly instructed the model to avoid human-centered assumptions, was more effective in shifting the LLM’s output toward a biological and ecological perspective. The dog-centric prompt yielded particularly interesting responses, with the LLM adopting a first-person canine perspective, however, performing as a human role playing. This strategy highlights the potential of language models to simulate or represent nonhuman viewpoints. However, these responses remained constrained by anthropocentric language patterns, underscoring the need for further exploration of new prompting strategies. For instance, combining elements of both the debiased and dog-centric prompts may foster richer, less human-biased outputs.

Nevertheless, this study has some limitations, which we would like to present and discuss to inform future research avenues upon this work. Our assessment of anthropocentrism was based on a subjective assessment rather than on a validated measurement. Additionally, we only analyzed a small subset of the collected data and, therefore, the analysis was focused on canine reactions to the stimulus of “food”. Furthermore, given the inherent human bias of language-based models like ChatGPT, future research should consider multimodal AI systems that incorporate non-verbal inputs and outputs, potentially enabling more authentic representations of nonhuman experiences.

Lastly, we would like to situate this exploration of anthropocentrism within broader paradigms centered on non-human entities. While this study frames the dog-centric perspective as distinct from, and at times in contrast to, an anthropocentric one, we acknowledge that these perspectives are not necessarily mutually exclusive. Humans and dogs share deep biological, social, and evolutionary ties, which can lead to overlaps in perception and behavior. Moreover, this work aligns with and advocates the more-than-human school of thought, which challenges the dominance of human-centered viewpoints and emphasizes the interconnectedness of humans and nonhumans. From this perspective, strict dichotomies between anthropocentric and nonhuman viewpoints may obscure the complexity of human–animal entanglements. Therefore, future research should move beyond assessing anthropocentrism as a single axis and instead examine other representations that reflect the co-existence and interdependence of humans and nonhumans.

References

[1] Clay Duncan and Ian Mcculloh. 2023. Unmasking Bias in Chat GPT Responses. In *Proceedings of the International Conference on Advances in Social Networks Analysis and Mining*. 687–691.

[2] Raluca Alexandra Fulgu and Valerio Capraro. 2024. Surprising gender biases in GPT. *Computers in Human Behavior Reports* (2024), 100533.

[3] Georgios P Georgiou. 2024. ChatGPT Exhibits Bias Toward Developed Countries Over Developing Ones, as Indicated by a Sentiment Analysis Approach. *Journal of Language and Social Psychology* (2024), 0261927X241298337.

[4] Sankalpa Ghose, Tse Yip Fai, Kasra Rasaee, Jeff Sebo, and Peter Singer. [n. d.]. The Case for Animal-Friendly LLMs. In *AAAI-2024 Workshop on Public Sector LLMs: Algorithmic and Sociotechnical Design*.

[5] Thilo Hagendorff, Leonie N Bossert, Yip Fai Tse, and Peter Singer. 2023. Speciesist bias in AI: how AI applications perpetuate discrimination and unfair outcomes against animals. *AI and Ethics* 3, 3 (2023), 717–734.

[6] Isabelle Hupont, Marina Wainer, Sam Nester, Sylvie Tissot, Lucía Iglesias-Blanco, and Sandra Baldassarri. 2024. Synocene, Beyond the Anthropocene: De-Anthropocentralising Human-Nature-AI Interaction. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*. 1–9.

[7] Jingling Li, Zeyu Tang, Xiaoyu Liu, Peter Spirtes, Kun Zhang, Liu Leqi, and Yang Liu. 2024. Steering LLMs Towards Unbiased Responses: A Causality-Guided Debiasing Framework. *arXiv preprint arXiv:2403.08743* (2024).

[8] Raphaël Millière and Charles Rathkopf. 2024. Anthropocentric bias and the possibility of artificial cognition. In *ICML 2024 Workshop on LLMs and Cognition*.

[9] Ben Mylius. 2018. Three types of anthropocentrism. *Environmental Philosophy* 15, 2 (2018), 159–194.

[10] Vandana Sharma, Naman Mishra, Vinay Kukreja, Ahmed Alkhayyat, and Ahmed A Elngar. 2023. Framework for Evaluating Ethics in AI. In *2023 International Conference on Innovative Data Communication Technologies and Application (ICIDCA)*. IEEE, 307–312.

[11] Peter Singer. 1981. *Animal liberation. Towards an end to man's inhumanity to animals.*

[12] Lindia Tjuatja, Valerie Chen, Tongshuang Wu, Ameet Talwalkwar, and Graham Neubig. 2024. Do llms exhibit human-like response biases? a case study in survey design. *Transactions of the Association for Computational Linguistics* 12 (2024), 1011–1026.

[13] Chunchen Xu and Xiao Ge. 2024. AI as a Child of Mother Earth: Regrounding Human-AI Interaction in Ecological Thinking. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*. 1–9.

Received 20 February 2007; revised 12 March 2009; accepted 5 June 2009