



Prompt Engineering for Hermeneutical Justice in LLMs

An Empirical Study on ADHD-Related Causal Reasoning

Sanjit Sankara Subramanian Lakshmi¹

Supervisor(s): Jie Yang¹, Anne Arzberger¹

¹EEMCS, Delft University of Technology, The Netherlands

A Thesis Submitted to EEMCS Faculty Delft University of Technology,
In Partial Fulfilment of the Requirements
For the Bachelor of Computer Science and Engineering
June 22, 2025

Name of the student: Sanjit Sankara Subramanian Lakshmi
Final project course: CSE3000 Research Project
Thesis committee: Jie Yang, Anne Arzberger, Myrthe Tielman
Email: S.SankaraSubramanianLakshmi@student.tudelft.nl

An electronic version of this thesis is available at <http://repository.tudelft.nl/>.

Abstract

Large Language Models are increasingly integrated into everyday applications, but their responses often reflect dominant cultural narratives, which can lead to misrepresentation of marginalized communities. This paper addresses the underexplored issue of hermeneutical epistemic injustice (HEI) in LLM outputs, particularly how these systems fail to accurately represent the lived experiences of people with ADHD when answering causal questions, and whether different prompting techniques can influence and improve the justice reflected in their responses. We introduce a practical framework for measuring HEI based on four proxies: intelligibility, conceptual fit, recognition of structural barriers, and expression style. Through a within-subjects user study with seven adults with ADHD, we evaluated three prompting strategies: Vanilla (baseline), Step-Back, and Human Persona + System 2. Our findings show that Human Persona + System 2 prompting stood out for its empathetic tone, balanced perspectives, and non-judgmental framing, thereby improving fairness across multiple HEI dimensions. Surprisingly, Vanilla prompts performed comparably well overall, while Step-Back responses offered clear practical information and contextual relevance, but were limited by an impassive, matter-of-fact tone. These results suggest that prompt design can meaningfully affect how well LLMs represent marginalized experiences. We conclude that advancing epistemic justice in generative AI requires thoughtful prompt design and may benefit from deeper engagement with affected communities to more accurately and respectfully represent their realities.

1 Introduction

Large Language Models (LLMs) are rapidly transforming our society, reshaping industries from healthcare and education to business and media [19]. As their influence grows across these domains, it becomes increasingly important to ensure the information people receive is representative and unbiased, especially given that these systems can reinforce societal biases, as their training data may be non-representative of the population or overly generalized [8]. Even when data is collected carefully, it can still carry historical and structural biases, leading to models that fail to account for the needs of marginalized social groups.

One form of bias that is especially relevant in this context is hermeneutical epistemic injustice (HEI). This occurs when individuals are unable to understand or express their own experiences due to a lack of shared social concepts or interpretive resources [7]. For example, a student with undiagnosed ADHD might struggle with focus, time management, and emotional regulation, but be labeled as lazy or careless in environments where neurodivergence is poorly understood, leaving them unable to articulate their difficulties in socially recognized terms. Such cases highlight the kind of epistemic harms that LLMs may replicate or even amplify in its outputs.

Dwivedi et al.[5] demonstrate that prompt design and in-context learning can mitigate gender bias in model outputs, reporting a 40% reduction in biased responses, across three distinct metrics that are based on the frequency of gendered associations and male-to-female pronoun ratios in generated LLM text. They further suggest that, although LLMs inevitably inherit biases from their training data, the extent to which these biases manifest in outputs can be significantly controlled through prompt design. Prompt engineering involves crafting natural language instructions to guide LLMs in generating relevant responses, without the need for retraining or fine-tuning the model [26]. Despite its potential, there remains a gap in understanding how prompt engineering can be leveraged to reduce hermeneutical epistemic injustices in LLM responses.

This research aims to address this gap by investigating whether prompts can be designed to elicit more hermeneutically just responses from LLMs. It focuses specifically on causal questions related to ADHD, which ask why certain outcomes occur. To explore this, we conducted a within-subjects user study in which participants with ADHD evaluated LLM-generated responses to causal questions. These types of questions often require a nuanced understanding of lived experiences that may be underrepresented or distorted in training data [8]. By focusing on this domain, the study seeks to contribute to the development of fairer, more inclusive and epistemically just generative AI systems for marginalized communities. The main research question this paper addresses is:

How can prompt design influence and improve the hermeneutical justice of LLM responses to causal questions about ADHD-related topics?

The remainder of this paper is organized as follows. Section 2 introduces the foundational concepts and literature that our work builds on. Section 3 outlines the methodology, explaining how the study is designed to produce reliable and repeatable results. Section 4 presents the experimental setup, and both the quantitative and qualitative results of the study. Section 5 reflects on the findings and interprets their implications. Section 6 discusses the ethical considerations of this research and addresses reproducibility. Finally, Section 7 summarizes the overall contributions of the study and suggests directions for future research in this domain.

2 Background

2.1 Hermeneutical Injustice in LLM Responses

Hermeneutical epistemic injustice (HEI) in LLMs occurs when these systems misrepresent or fail to capture the experiences of marginalized communities due to gaps in cultural or contextual understanding [14]. Kay et al. [14] introduce the concept of *generative hermeneutical ignorance* to describe how language models, while appearing knowledgeable, often reproduce dominant narratives and overlook key aspects of lived experience, especially those that fall outside mainstream representation.

Generative hermeneutical ignorance differs from traditional HEI by stemming directly from AI models' inherent technological limitations rather than gaps in shared conceptual resources. Kay et al. [14] argue that these systems "lack the embodied and cultural knowledge humans acquire through lived experience", leading to misinterpretation of marginalized voices even without discriminatory training data. They further argue that beyond perpetuating existing biases, these systems can also produce entirely new forms of misinterpretation. As generative models become more integrated into everyday tools, their outputs can shape how society understands marginalized identities, potentially causing real harm to individuals and communities, thus making it vital to address hermeneutical justice in LLM responses.

2.2 Understanding ADHD as a Lived and Misunderstood Experience

Attention-deficit/hyperactivity disorder (ADHD) is a common neurodevelopmental condition affecting up to 7% of adults worldwide, that is characterized by patterns of inattention, hyperactivity, and impulsivity [25] [32]. For many people with ADHD, the condition leads to profound social and emotional consequences that are frequently misunderstood or dismissed by others.

In practice, ADHD can lead to serious social impairments. People with the condition often struggle with tasks like timekeeping, organization, or staying in touch, yet these behaviors are commonly interpreted by others as laziness, carelessness, or lack of empathy, a disconnect highlighted by Ginopp et al. [9]. As one participant from their study reflected, "Even friendships... people aren't realizing that me being late isn't a reflection on how much I care about them."

Miscommunications are also common, particularly when neurotypical listeners misinterpret ADHD-related behaviors as rudeness or inattention. In the same study, another individual described their experience: "When I meet people, you can kind of see the wheels turning... They're thinking about what you said, and it seems like they didn't take it like how you meant it... And then they think you're weird because you're over explaining everything" [9]. These examples reveal the disconnect between lived experiences and the dominant interpretive frameworks used to understand them.

Given ADHD's emotional toll and frequent misrepresentation, it serves as an especially relevant lens through which to investigate representational harms in AI systems. If LLMs inherit society's interpretive blind spots, then studying how they respond to ADHD-related questions offers a meaningful opportunity to surface and potentially reduce hermeneutical epistemic injustice.

2.3 Why Focus on Causal Questions?

Causal questions are particularly valuable for potentially uncovering HEI in LLM outputs. Unlike descriptive questions, which ask for surface-level facts (e.g., "What is ADHD?"), causal questions require the model to reason about how one experience leads to a particular outcome, such as asking "How does untreated ADHD in childhood affect social relationships later in life?" - which is a question later employed in the user study (see section 4.1) [12]. This makes them more likely to expose where the model lacks interpretive resources or misrepresents marginalized experiences.

Evidence from BiasCause, a recent benchmark study, highlights this vulnerability [30]. The authors show that across four leading language models, performance on socially biased causal reasoning tasks remains extremely poor, often below 36% accuracy. These results are consistent with broader findings on social bias, where language models exhibit higher error rates when reasoning about age and disability, compared to attributes like race [30]. Since HEI falls under the broader category of social bias, and ADHD is a disability, causal questions about ADHD-related topics are a productive context for observing and potentially correcting such injustices.

3 Methodology

3.1 Research Design

This study employs a mixed-method experimental design to investigate whether prompt design can improve the hermeneutical justice of LLM responses to causal questions about ADHD-related topics. The investigation is guided by three sub-research questions:

RQ1: What measurable proxies can be used to assess HEI in LLM responses?

RQ2: Which prompting techniques have the potential to reduce HEI in responses to causal questions?

RQ3: To what extent do the prompt techniques improve the hermeneutical epistemic justice in LLM responses to causal questions about ADHD-related topics?

RQ1 and RQ2 are addressed through literature review. RQ1 helps define the evaluation criteria by identifying measurable proxies of HEI, while RQ2 guides the selection of prompt techniques used in the experiment. RQ3 is explored through prompt engineering and a user study with participants who have ADHD, aiming to evaluate whether different prompting techniques improve the hermeneutical justice of LLM responses.

3.2 Identifying Proxies of Hermeneutical Epistemic Injustice (RQ1)

Hermeneutical epistemic injustice is a broad and abstract construct. It is difficult to assess directly and Fricker’s original account provides philosophical grounding but not a fixed framework for evaluation. To make the construct measurable for this study, HEI is defined in a more tangible manner, through four proxies identified through literature review and simplified for participants in the survey, as common in scientific research.

To strengthen the trustworthiness of the proxies developed in this study, inspiration was drawn from Clark and Watson’s guidance on construct validation, especially their emphasis on clearly defining the construct (see Appendix F) and understanding it within its full theoretical context through a focused literature review [4]. In line with this, key aspects of HEI, as described in Fricker’s seminal work and related research, were examined closely. Fricker’s account, widely regarded as the foundation of work on epistemic injustice, served as a strong conceptual starting point for grounding both the construct and the proxies.

Intelligibility. This attribute assesses whether a response makes the experience being described clear and understandable. It captures how well the LLM explains the topic in a way that makes sense - not so simple that it overlooks important details, and not so complex that it becomes difficult to follow. Drawing on Fricker’s framing, intelligibility concerns whether the model attempts to bridge interpretive gaps instead of blaming the speaker for miscommunication [7].

Conceptual Fit. Conceptual fit evaluates how well the ideas and examples used in the response align with the lived experiences of people with ADHD. It considers whether the LLM avoids stereotypes, uses accurate language, and reflects common challenges described in research or first-hand accounts. As Kay et al. [14] point out, models may appear fluent but still fail to capture the grounded realities of marginalized groups due to a lack of cultural and embodied knowledge. Conceptual fit ensures that responses feel contextually true and relevant to the person they aim to represent.

Recognition of Structural Barriers. This attribute captures whether the response acknowledges broader social or institutional factors that shape the lived experience, such as inflexible school systems, inaccessible workplaces, or social stigma. HEI is often tied to structural inequalities, which prevent some experiences from being widely understood [1]. A response that only focuses on individual challenges, without naming systemic barriers, risks reinforcing an unjust narrative.

Expression Style. Even when content is accurate and understandable, a response can still be unjust if the tone is dismissive, condescending, or overly clinical. This attribute assesses whether the LLM responds with empathy, care, and inclusive language. Fricker notes that hermeneutical injustice may emerge not just from what is said, but how it is said [7].

Together, these four proxies offer a practical lens for identifying hermeneutical injustice in LLM responses. They allow for consistent evaluation across different participants and

questions, while staying grounded in the theory of hermeneutical justice.

3.3 Selecting Prompting Techniques (RQ2)

Improving the reasoning capabilities of LLMs has been shown to reduce bias and promote fairness in their responses [11]. Since HEI is a form of bias rooted in misrepresentation and a lack of interpretive resources, prompting techniques that encourage clearer reasoning or better contextual understanding may play a role in reducing HEI. This study therefore sought to identify and evaluate prompting techniques with demonstrated potential for bias reduction.

To survey existing options, this research began with a review of The Prompt Report, which catalogs over 58 prompting techniques across different use cases [23]. Given scope and time limitations, a targeted review was conducted focusing on techniques with evidence of improving reasoning or reducing social bias. Techniques were shortlisted if they demonstrated either **(a) empirical improvements in reasoning quality** or **(b) proof of reductions in harmful bias**, as both characteristics serve as useful proxies for potential reductions in HEI.

Six techniques were initially reviewed: Chain-of-Thought (CoT), which encourages intermediate reasoning steps and improves performance on reasoning tasks [28]; Re-reading (RE2), which enhances comprehension through double-pass processing [31]; Step-Back Prompt, which promotes contextual awareness by posing higher-level questions first [33]; Chain-of-Thought Factored Decomposition, which breaks complex questions into manageable components [6]; Self-Consistency, which samples multiple reasoning paths for internal validation [27]; and Human Persona + System 2, which combines deliberate reasoning with a human perspective [13]. A more detailed breakdown of the techniques reviewed and the reasons for their initial selection has been added to Appendix G.

From this pool, **Step-Back** and **Human Persona + System 2** were selected for inclusion in the experiment based on their stronger conceptual alignment with the four proxies of HEI. While other techniques demonstrated improvements in general reasoning or bias mitigation, these two were more directly relevant to the interpretive and representational challenges at the core of HEI.

Step-Back encourages the model to broaden its contextual scope before answering specific questions. Prior research shows that prompting with a higher-level, related question before the target question can improve the model’s ability to retrieve relevant background knowledge and generate more nuanced responses [33]. For example, before asking "How does undiagnosed ADHD in adulthood affect a person’s ability to maintain healthy relationships?", a step-back question might be "What are some common reasons why people struggle to maintain healthy relationships over time?" This capacity to elicit broader contextual reasoning aligns closely with the proxies of Conceptual Fit and Recognition of Structural Barriers.

Human Persona + System 2, which emphasizes deliberate reasoning framed through human perspectives, supports the proxies of Intelligibility and Expression Style. Kamruzzaman et al. [13] found that this approach significantly reduced stereotypical judgments by 19%, further supporting its relevance for mitigating representational harms. A typical prompt used to implement this technique might be: "Adopt the identity of a person who answers questions slowly and thoughtfully. Their answers are effortful and reliable. [Question]".

Due to time constraints and the need to minimize participant burden, only a limited number of techniques could be evaluated. These two were therefore prioritized for their

conceptual relevance to HEI. Future work would benefit from evaluating all shortlisted techniques to allow for a more comprehensive comparison.

3.4 User Study and Evaluation (RQ3)

This subsection details how the user study was designed to evaluate the impact of different prompting techniques on the perceived hermeneutical justice of LLM responses.

Recruitment. Seven adult participants who self-identified as having ADHD were recruited through community outreach and personal networks. Eligibility required participants to have lived experience with ADHD. All participants provided informed consent prior to participation and the study protocol was approved by the TU Delft Human Research Ethics Committee (HREC).

Model Access and Configuration. The study used the GPT-4o model via OpenAI’s ChatGPT web interface (<https://chat.openai.com>) to generate responses, aligning with the version accessed by real-world users. This approach was chosen over running a local model or using the API in order to accurately assess the behavior of the LLM as it would be experienced by people with ADHD in typical, everyday use.

Question Design. Participants were assigned to evaluate model-generated responses to one of three causal questions concerning ADHD in the domains of education, employment, and healthcare. These domains were selected due to their relevance in daily life and their potential to reveal HEI. The causal questions were generated using GPT-4o by explicitly prompting it to produce questions likely to elicit HEI in responses, based on prior literature identifying the social challenges associated with ADHD, specifically the research by Ginopp et al. [9]. The questions used in the study are mentioned in Section 4.1.

Prompting Conditions. For each question, three prompt techniques were used: Vanilla (baseline), Step-Back, and Human Persona + System 2. Three responses were generated for each technique to account for the inherent non-determinism of LLM outputs. To avoid contextual influence from prior interactions, each LLM response was generated in a new thread to ensure that responses were based solely on the prompt provided, without any memory from previous inputs [18].

Experimental Design and Sampling. We employed a within-subjects design to assess how prompting techniques influenced perceived hermeneutical justice. This design was chosen because having each participant evaluate all prompting techniques reduces the impact of individual differences, allowing for a more sensitive comparison than a between-subjects design where different participants evaluate different techniques.

Each participant rated one LLM response per prompting technique for a single ADHD-related causal question. To mitigate topic-specific biases, participants were randomly assigned to one of three distinct ADHD topics. The prompting technique served as the sole fixed effect in our analysis, while topic variability was treated as a random effect to support generalization across contexts. This approach aligns with similar within-subjects designs generalizing across stimulus sets [29].

Randomization. To reduce bias from selectively choosing favorable responses, one output per technique was randomly selected for each participant. Randomization ensured that each participant engaged with a unique set of outputs, making it more likely that differences in ratings were due to the prompting technique itself rather than the specific responses shown [24]. This strengthens internal validity, meaning the study is more trustworthy in attributing differences in participant ratings to the prompting techniques.

Evaluation Procedure. Participants rated each response using a 5-point Likert scale

for four proxies of hermeneutical epistemic justice: Intelligibility, Conceptual Fit, Recognition of Structural Barriers, and Expression Style. Clear and concise definitions for each of the four proxies were provided to participants within the survey (see Appendix A). To complement these numeric assessments, participants also answered four open-ended questions (Appendix E), each corresponding to one proxy. These open responses provided insight into how participants perceived the nuanced presence or absence of justice-related elements in each response.

Quantitative Analysis. Likert-scale ratings were aggregated and compared across prompting techniques and proxies. This allowed for a structured comparison of how effectively each technique supported the recognition and articulation of ADHD-related experiences, revealing which methods most consistently supported hermeneutical justice.

Qualitative Analysis. The open-ended responses were analyzed using an approach inspired by Reflexive Thematic Analysis (RTA), drawing on Braun and Clarke’s work [2]. While not a strict application of RTA, this approach was chosen for its flexibility and focus on interpreting meaning across participants’ expressed experiences, while guided by the proxies created through literature review. Unlike content analysis, which focuses on counting and categorizing explicit terms or ideas, this RTA-inspired method allows for the identification of broader patterns and themes within participant responses. In contrast to discourse analysis, which examines how language is used to produce certain effects, this approach places greater emphasis on what participants are trying to communicate.

This study used a mixed analytical approach. The overall structure of the analysis was **deductive**, guided by the four predefined proxies of HEI, which served as the main themes. Alongside this, the sub-themes were identified **inductively**, based on patterns that emerged directly from participants’ feedback. Once identified, these sub-themes were grouped under the HEI proxy they most closely related to, allowing the analysis to stay grounded in theory while also reflecting the voices of participants.

3.5 Positionality Statement

We recognize that our personal experiences influences how this study was designed and interpreted. The main researcher is a Bachelor’s student in Computer Science and Engineering, with some prior knowledge in Human-Computer Interaction and a growing interest in the ethical dimensions of AI. While we do not identify as part of the ADHD community, we acknowledge that large language models may be trained on dominant cultural narratives that may misrepresent or exclude marginalized perspectives. This research was approached with the aim of improving hermeneutical justice in LLM responses. We conducted a thematic analysis using predefined proxies of hermeneutical epistemic injustice as our analytic lens, remaining mindful of our outsider position and allowing participants’ perspectives to guide interpretation by closely attending to their evaluative terms, how these related to the attributes, and avoiding overgeneralization beyond what was expressed.

4 Experimental Setup and Results

4.1 Experimental Setup

Each participant was randomly assigned to one of three causal questions, resulting in three participants evaluating Question A, and two participants each evaluating Questions B and C. For each causal question, a separate survey was created containing three GPT-4o generated

responses, one per each prompting technique. Although the ADHD topic varied between participants, the core comparison across prompting strategies remained constant, reflecting the within-subjects design.

The three causal questions used in the user study are listed below, while the prompt used to generate them and the rationale for their selection are detailed in Appendix B.

Question A: How does untreated ADHD in childhood affect social relationships later in life?

Question B: Why do many adults with ADHD struggle with maintaining employment stability?

Question C: Why do people with ADHD often report feeling misunderstood in health-care settings?

Each participant evaluated only one causal question to limit cognitive load and allow for meaningful reflection. Participants were allocated approximately 10 minutes to complete their evaluation. The study was conducted online using a survey. While participants completed the survey independently, we were present on standby in a Microsoft Teams call in case any clarifications were needed during the session. Before beginning the survey, participants were given a short verbal introduction explaining the evaluation task and how to complete the survey.

4.2 Quantitative Analysis

This section presents the quantitative results of the user study, which assessed LLM-generated responses across three prompting techniques using Likert scale ratings. Because Likert items represent ordinal data, we report mode and median scores instead of parametric statistics like mean and standard deviation, as recommended in existing survey methodology literature [22]. To enhance interpretability and allow effective visual comparisons between techniques, we use diverging stacked bar charts. These charts are widely recommended for Likert-scale data due to their clarity and ability to support cross-category comparison [20].

Each of the four diverging stacked bar charts (Figure 1) corresponds to one of the four proxies of hermeneutical epistemic injustice: Intelligibility, Conceptual Fit, Recognition of Structural Barriers, and Expression Style. The vertical axis lists the three prompting techniques used in the study: Vanilla, Step-Back, and Human Persona + System 2. The horizontal axis shows the percentage of participant ratings across the five-point Likert scale. These charts answer the question, "How well does the response capture [proxy]?" by visualizing the distribution of participant ratings per technique.

Alongside the visualizations, three summary tables present the mode and median for each prompting technique across the four proxies. The median indicates the central or "typical" response, helping to summarize where participant opinion tends to cluster. The mode, in contrast, shows the most frequently chosen score, capturing where consensus was strongest. The Likert scale used in the survey ranges from 1 to 5, where 1 = Not at all, 2 = Slightly, 3 = Somewhat, 4 = Mostly, and 5 = Completely. Higher scores indicate stronger agreement that the response captured the relevant proxy of hermeneutical justice.

Table 1 shows that for the Step-Back prompting technique, the median and mode both equal 4 for Intelligibility, Conceptual Fit, and Expression Style. This indicates a strong consensus that the technique consistently performed well on these aspects.

Table 2 indicates particularly high ratings for Expression Style, where both mode and median equal 5. Across the other proxies, mode and median values mostly align at 4, pointing to a stable and generally favorable impression of the technique.

Table 3 shows that for the Vanilla prompt, the median and mode are both 4 for Intelligibility, Conceptual Fit, and Expression Style, suggesting moderate consistency. However, Recognition of Structural Barriers shows a median of 2 with mode values at 1 and 2, revealing a cluster of lower ratings.

Together, the charts and tables offer a descriptive overview of how each prompting technique was received in relation to each proxy of hermeneutical justice.

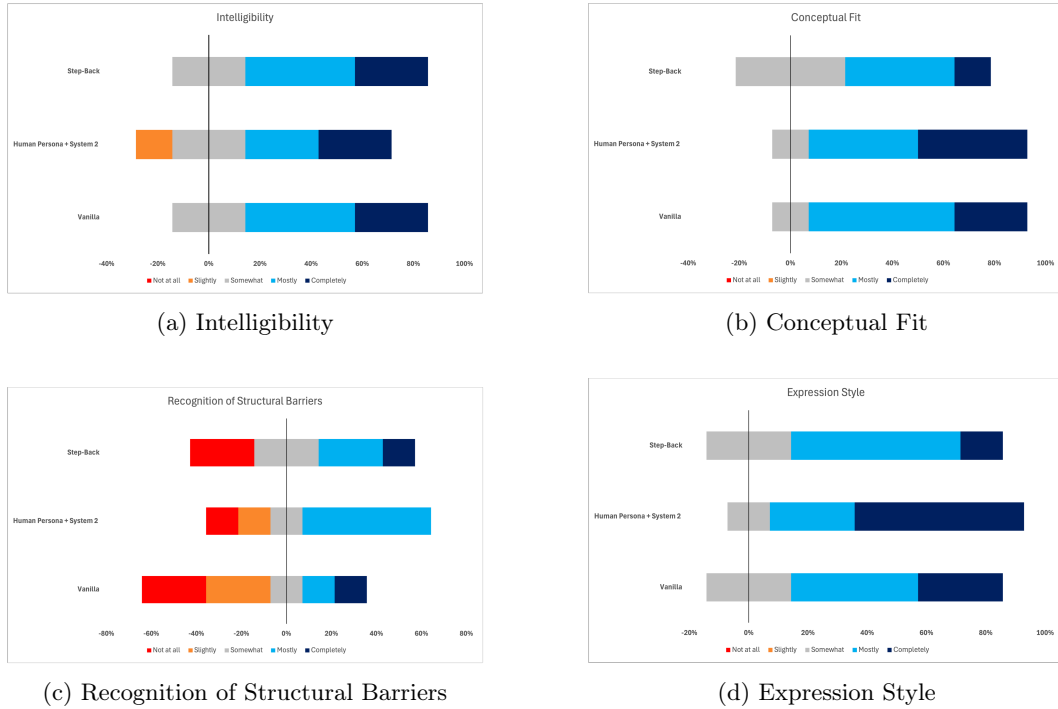


Figure 1: Diverging bar charts illustrating different proxies of HEI across multiple prompting techniques.

	Intelligibility	Conceptual Fit	Recog. Of Structural Barriers	Expression Style
Mode	4	3,4	1,3,4	4
Median	4	4	3	4

Table 1: Mode and Median Ratings for the Step-Back Prompting Technique Across HEI Proxies

	Intelligibility	Conceptual Fit	Recog. Of Structural Barriers	Expression Style
Mode	3,4,5	4,5	4	5
Median	4	4	4	5

Table 2: Mode and Median Ratings for the Human Persona + System 2 Prompting Technique Across HEI Proxies

	Intelligibility	Conceptual Fit	Recog. Of Structural Barriers	Expression Style
Mode	4	4	1,2	4
Median	4	4	2	4

Table 3: Mode and Median Ratings for the Vanilla Prompting Technique Across HEI Proxies

4.3 Qualitative Analysis

We conducted a thematic analysis inspired by Braun and Clarke’s RTA [2], combining deductive coding based on the four predefined proxies of HEI with inductive coding to identify sub-themes within each proxy. Participants’ open-ended feedback was analyzed and organized under the themes of Intelligibility, Conceptual Fit, Recognition of Structural Barriers, and Expression Style. In this section, participant quotes are attributed using identifiers (P1 = Participant 1, P2 = Participant 2, etc.) along with the specific question (A, B, or C) they were responding to. The detailed codebook with definitions and examples of all themes and sub-themes can be found in Appendix H.

Intelligibility. Participants emphasized the importance of clarity and structure in the LLM responses. Structural clarity was a key concern, with P4 noting for Question B, "All of the responses were a big block of text; splitting it into 2 paragraphs made it easier to separate concepts and keep track of where I am in the text." P2 found the Step-Back response for Question A "a little confusing because it’s mostly only explaining how ADHD affects adults, barely explaining the effect of untreated ADHD in childhood," highlighting the need for developmental focus across life stages. Language accessibility also varied, with P7 pointing out that the Human Persona + System 2 response for Question C "was least intelligible, the words it used to describe certain symptoms, consequences or patterns were not helpful of my understanding of them," demonstrating challenges in making complex information intelligible without losing important details.

However, P5 specifically praised how the Human Persona + System 2 response for Question B "puts the reader in the shoes of both people with ADHD and the ones of clinicians/others, which helps outlining the issues people with ADHD face and why it happens, without putting the 'blame' on anyone specifically," suggesting that intelligibility also depends on presenting multiple perspectives in a way that clarifies the root of misunderstanding without oversimplification.

Conceptual Fit. Feedback revealed nuanced perspectives on how well responses reflected the lived experience of ADHD. ADHD subtype representation was crucial, with P1 explaining for Question A, "For me, I had a very different type of ADHD, in which impulsivity is not present but only inattention (ADHD-I). Therefore I did not regard those materials talking exclusively about me... this need to be properly addressed." P5 noted for Question B that all responses "lack an emphasis on how ADHD manifests itself widely differently in different people... It feels like all of them are slightly stereotypic in descriptions, through using big words and terms such as 'emotional dysregulation' but not emphasizing how this has wide manifestations." Recognition of strengths was also missing, as P3 observed for Question A that "all could add extra information about the positive things people go through despite all the negatives." These comments indicate that conceptual fit depends on balancing accurate clinical detail with diverse and authentic lived experiences.

Recognition of Structural Barriers. Participants generally observed limited inclusion of systemic factors in the responses. P1 noted for Question A, "All responses have little mention of the social settings, except that response 3 [Step-Back] mentioned challenges in

work environment. I think school should definitely be mentioned as it's an important part of childhood and can really affect children with untreated ADHD." P3 remarked for Question A that responses "focused more on the individual rather than external factors that are hard or nearly impossible for someone with ADHD to 'fix' or control." Similarly, P7 highlighted for Question C, "The misunderstanding of ADHD in children, which was especially more prevalent when these adults were young, also plays a big part," emphasizing educational system impact. These reflections point to a need for responses that contextualize ADHD within broader social and institutional frameworks rather than focusing on individual deficits.

Expression Style. Tone and empathy were central to participant evaluations. Clinical versus personal tone was a key concern, with P3 describing the Vanilla and Step-Back responses for Question A as sounding "clinical," while P1 noted that all responses "feel very academic - a tone that suggests the content to be persuasive and scientific, though I have to read with reasoning actively to differentiate if it is really logical." In contrast, P3 said for Question A, "Response 2 [Human Persona + System 2] felt more empathetic as it didn't give absolutes," and P5 noted for Question B, "2nd [Human Persona + System 2] was the most empathetic and it is quite visibly different in tone from the others." Conversely, P7 found the Step-Back response for Question C problematic, saying it "felt like it was written by someone who googled what ADHD is for an hour, and then tried their best with that," indicating a perceived lack of genuine understanding. Empathy and inclusivity in expression were thus crucial to participants' sense of hermeneutical justice.

These verbatim illustrate where LLM responses meet or fall short of the four proxies of HEI, highlighting the importance of balancing clear, conceptually accurate content with empathetic tone and recognition of structural challenges.

5 Discussion

This study investigated how prompt design can influence and improve the hermeneutical justice of LLM responses to causal questions about ADHD-related topics. The findings demonstrate that certain prompt techniques can enhance hermeneutical epistemic justice, though important limitations remain.

These results contribute to an emerging body of research that examines epistemic injustice in AI systems. Recent work has identified hermeneutical injustice as a critical concern in generative AI, where marginalized groups are misrepresented in AI-generated knowledge [14]. Specifically regarding neurodivergent experiences, emerging research examining LLM usage by neurodivergent individuals has identified key challenges in AI-generated responses, including lack of personal voice and neurotypical biases that fail to authentically represent diverse neurodivergent perspectives [3]. Addressing these issues, the present study explores prompt engineering techniques aimed at improving hermeneutical justice in LLM responses for individuals with ADHD.

Key Findings and Significance. The Human Persona + System 2 prompting technique emerged as the most effective approach, achieving the highest ratings for Expression Style (mode and median of 5) and consistently strong performance across other proxies. Compared to the Vanilla prompting (baseline), this technique showed clear improvement in Recognition of Structural Barriers (median increased from 2 to 4) and Expression Style (median increased from 4 to 5). This finding aligns with research on Human Persona + System 2 prompting technique, that illustrates reductions in social bias in LLMs [13].

Surprisingly, Vanilla prompting performed comparably well in three proxies (Intelligibility, Conceptual Fit, and Expression Style), challenging more complex prompting tech-

niques. This finding aligns with recent research showing that advanced models like GPT-4 can achieve high accuracy in medical diagnoses and exhibit reduced bias across race and ethnicity [10]. However, all techniques struggled with Recognition of Structural Barriers, where even the best-performing approach failed to adequately address systemic factors affecting ADHD experiences.

Supporting Evidence. The quantitative data shows clear differentiation between techniques, with Human Persona + System 2 consistently outperforming others in most proxies. The qualitative analysis reveals why this approach was successful. Participants distinguished between "clinical" and "empathetic" tones, finding persona-based responses more authentic and grounded in real experience. Crucially, participants valued empathy and genuine understanding over responses that sounded "like it was googled for an hour," while still appreciating accurate clinical terminology when it enhanced conceptual fit and intelligibility. This balance between empathetic tone and technical accuracy appears central to achieving hermeneutical justice.

Addressing the Research Gap. This research directly addresses the gap in understanding how prompt design can influence and improve hermeneutical justice in LLM responses about ADHD experiences. The results provide empirical evidence that prompt engineering can serve as a method for reducing HEL, particularly through human perspective-taking and deliberative reasoning processes. This finding extends existing theoretical frameworks that define hermeneutical injustice as barriers to accessing meaningful information as an epistemic agent [17], by demonstrating practical interventions that can mitigate such barriers in AI systems.

Critical Limitations. All prompting techniques inadequately addressed *Recognition of Structural Barriers* proxy. Participants consistently noted absent educational system impacts, and other systemic factors shaping ADHD experiences. This limitation aligns with broader critiques of knowledge-enhanced language models, which argue that epistemic injustices persist because marginalized groups remain underrepresented in knowledge production and collective AI resources [15]. Additionally, responses underrepresented ADHD diversity, lacked strength-based perspectives and displayed stereotypical descriptions. These findings echo concerns about algorithmic profiling perpetuating hermeneutical injustice by reinforcing dominant narratives while marginalizing alternative perspectives [16].

Relationships Among Findings and Looking Ahead. The study reveals important trade-offs between different aspects of hermeneutical justice. While Human Persona + System 2 excelled in empathy, some found its responses less intelligible due to complex terminology. This suggests optimal hermeneutical justice may require hybrid approaches combining Step-Back's clarity with Human Persona + System 2's empathetic representation.

6 Responsible Research

This section outlines the ethical and broader social responsibilities considered throughout the study. It addresses how the research was conducted with integrity, ensured participant privacy, and remained attentive to potential limitations, biases, and the broader impact of working with marginalized communities in the context of generative AI.

Reproducibility and Participant Context. While this study is reproducible in principle, the small sample size (7 participants) limits generalizability. All participants were students or have just graduated, aged between 18-26, including both men and women from various nationalities. The question on employment stability may have had limited relevance for participants without direct work experience. Future replications could benefit from

recruiting more demographically and experientially diverse participants, and from tailoring causal questions more closely to their lived contexts. To help future researchers understand the interpretive lens used in this study, Section 3.5 outlines the researcher’s background and potential biases during analysis, through a positionality statement.

Ethics, Privacy, and Data Protection. This study involved human participants and adhered to code of conduct set by the faculty. Participation was entirely voluntary, and informed consent was obtained. Participants were made aware of the potentially sensitive nature of the questions, particularly regarding past ADHD-related experiences that may include relating to traumatic experiences in their lives. No personal identifiers, IP addresses, or sensitive open-ended data were recorded. Responses were anonymized and stored securely, used solely for academic purposes. Given that some questions touch on emotionally difficult topics, care was taken to allow participants to exit the study at any time without consequence.

In line with TU Delft’s Code of Conduct, this study was conducted with a strong commitment to the university’s core values of **Respect** and **Trust** [21]. Respect was upheld by ensuring the survey was designed with inclusive, non-judgmental language and by encouraging participants to share perspectives that reflected their lived experiences with ADHD, free from stigma or stereotyping. We approached participant input with an open mind, consistent with TU Delft’s emphasis on respectful engagement and non-discrimination. Trust was maintained through informed consent, anonymized data collection, and transparency about the study’s purpose. By taking responsibility for ethical conduct throughout, we aimed to uphold both personal and institutional integrity and ensure the trust of participants and the wider research community.

Research Integrity and Transparency. The entire study design, including survey questions, response generation methods, and analysis framework, was documented in detail to support transparency and replicability. The prompting techniques used were selected through literature review (see Section 3.3), and the specific responses evaluated by participants are included in the Appendix D. The exact prompts used to generate the responses are also included in Appendix C for reference.

To support clear communication, ChatGPT was used during the writing phase of this research to refine grammar, improve sentence flow, and enhance clarity. For instance, prompts such as "Refine the following for impact, clarity, and flow. Ensure there are no grammatical errors: [phrase/sentence/paragraph]" were used throughout the writing process. No content was generated solely by ChatGPT; its role was limited to refining the clarity of ideas and supporting language improvements.

Broader Impacts and Future Considerations. As generative AI becomes increasingly embedded in domains such as education, healthcare, and public services, ensuring that these systems fairly represent marginalized groups becomes a matter of responsibility for the researchers and engineers in this domain. The performative nature of LLM responses can shape perceptions of people with ADHD, especially among users with limited prior understanding. By improving how these models respond to causal questions related to ADHD, we aim to contribute to more equitable and fairer AI systems. Future work should include direct involvement from affected communities in co-design to further enhance fairness and interpretive alignment.

Limitations. While this study provides initial insights into how prompt design can influence HEI in LLM responses in the context of ADHD, several limitations constrain the scope and generalizability of the findings:

- The study focused solely on a single LLM (GPT-4o), and results may differ with other

models.

- Relying on LLM outputs introduces potential bias from the model’s training data, which may have influenced the phrasing of the generated causal questions used in the user study.
- The study tested only zero-shot prompting techniques, leaving out few-shot and multi-shot prompting methods that may offer different advantages in generating just responses.
- Although six promising prompting techniques were reviewed, only two were selected for experimentation due to time and feasibility constraints, limiting broader comparative analysis.
- The small sample size (seven participants) and narrow demographic profile (young adults) limit the generalizability of findings to the broader ADHD population.
- The four proxies of HEI were developed through a theory-driven approach, rather than being co-designed with members of the ADHD community, which may limit their relevance to lived experiences.
- Topic randomization (assigning different ADHD-related questions to participants) supports generalizability, but future work should explicitly investigate topic-specific effects to better understand their influence.
- No external validation (e.g. inter-rater reliability) was conducted on the thematic analysis, which could have strengthened the trustworthiness of qualitative findings.

7 Conclusions and Future Work

This study proposes a practical approach for measuring HEI in LLMs, identifies promising prompt design strategies, and evaluates their impact through a within-subjects user study with seven adults with ADHD. Human Persona + System 2 prompting consistently leads to more hermeneutically just responses, particularly through its empathetic tone, non-judgmental framing, and inclusion of diverse perspectives. Interestingly, the Vanilla prompt performs comparably well for several proxies, challenging assumptions about the necessity of complex prompting. Step-Back prompts offer strong contextual clarity, though their impassive tone limits perceived empathy.

Future work could also revisit all reviewed prompting techniques, including Chain-of-Thought, Re-reading, Factored Decomposition, and Self-Consistency, to further explore their potential for improving hermeneutical justice. To improve statistical power and generalizability, future studies should recruit a larger, more diverse participant pool and apply inferential statistical analyses in addition to the descriptive summaries used here due to time and sample constraints. Expanding the scope to include other forms of neurodivergence and adopting co-design approaches with affected communities could further strengthen the interpretive validity of both the HEI proxies and the evaluated prompting strategies. As a final consideration, experimenting with temperature settings may help clarify the role of creativity in shaping epistemic justice in responses.

References

- [1] Amiel Bernal. The epistemic injustice anthology: A review of the routledge handbook of epistemic injustice. *Social Epistemology Review and Reply Collective*, 6(11):1–8, 2017.
- [2] Virginia Braun and Victoria Clarke. *Thematic Analysis: A Practical Guide*. Sage Publications, 2021.
- [3] Buse Carik, Kaike Ping, Xiaohan Ding, and Eugenia H. Rho. Exploring large language models through a neurodivergent lens: Use, challenges, community-driven workarounds, and concerns, 2024.
- [4] L. A. Clark and D. Watson. Constructing validity: New developments in creating objective measuring instruments. *Psychological Assessment*, 31(12):1412–1427, 2019.
- [5] Shivam Dwivedi, Sudeshna Ghosh, and Shweta Dwivedi. Breaking the bias: Gender fairness in llms using prompt engineering and in-context learning. *Rupkatha Journal on Interdisciplinary Studies in Humanities*, 15(4), 2023.
- [6] Oluwole Fagbohun, Rachel M. Harrison, and Anton Dereventsov. An empirical categorization of prompting techniques for large language models: A practitioner’s guide. *arXiv preprint arXiv:2402.14837*, 2024.
- [7] Miranda Fricker. *Epistemic Injustice: Power and the Ethics of Knowing*. Oxford University Press, 2007.
- [8] Isabel O. Gallegos, Ryan A. Rossi, Joe Barrow, Md Mehrab Tanjim, Sunghul Kim, Franck Dernocourt, Tong Yu, Ruiyi Zhang, and Nesreen K. Ahmed. Bias and fairness in large language models: A survey. *Computational Linguistics*, 50(3):1097–1179, 2024.
- [9] Collie M. Ginopp, Norman R. Greenberg, Grace Macdonald-Gagnon, Gustavo A. Angarita, Krysten W. Bold, and Marc N. Potenza. The experiences of adults with adhd in interpersonal relationships and online communities: A qualitative study. *SSM - Qualitative Research in Health*, 3:100223, June 2023.
- [10] N Ito, S Kadomatsu, M Fujisawa, K Fukaguchi, R Ishizawa, N Kanda, D Kasugai, M Nakajima, T Goto, and Y Tsugawa. The accuracy and potential racial and ethnic biases of gpt-4 in the diagnosis and triage of health conditions: Evaluation study. *JMIR Medical Education*, 9:e47532, 2023.
- [11] Sanchit Kabra, Akshita Jha, and Chandan K. Reddy. Reasoning towards fairness: Mitigating bias in language models through reasoning-guided fine-tuning. *arXiv preprint arXiv:2504.05632*, 2025.
- [12] Steven J. Kamper. Types of research questions: Descriptive, predictive, or causal. *Journal of Orthopaedic & Sports Physical Therapy*, 50(8):468–469, 2020.
- [13] Mahammed Kamruzzaman and Gene Louis Kim. Prompting techniques for reducing social bias in llms through system 1 and system 2 cognitive processes. 2024.
- [14] Jenny Kay, Atoosa Kasirzadeh, and Shakir Mohamed. Epistemic injustice in generative ai. In *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, volume 7, pages 684–697, 2024.

- [15] Angelie Kraft and Eloïse Soulier. Knowledge-enhanced language models are not bias-proof: Situated knowledge and epistemic injustice in ai. pages 1433–1445. Association for Computing Machinery, 2024.
- [16] S. Milano and C. Prunkl. Algorithmic profiling as a source of hermeneutical injustice. *Philosophical Studies*, 182:185–203, 2025.
- [17] Martin Miragoli. Conformism, ignorance injustice: Ai as a tool of epistemic oppression. *Episteme*, pages 1–19, 2024.
- [18] Valentina Porcu. The role of memory in llms: Persistent context for smarter conversations. *International Journal of Scientific Research and Management (IJSRM)*, 12(11):1673–1691, 2024.
- [19] Mohaimenul Azam Khan Raiaan, Md. Saddam Hossain Mukta, Kaniz Fatema, Nur Mohammad Fahad, Sadman Sakib, Most Marufatul Jannat Mim, Jubaer Ahmad, Mohammed Eunus Ali, and Sami Azam. A review on large language models: Architectures, applications, taxonomies, open issues and challenges. *IEEE Access*, 12, 2024.
- [20] Naomi B. Robbins and Richard M. Heiberger. Plotting likert and other rating scales. In *Section on Survey Research Methods – JSM 2011*. American Statistical Association, 2011.
- [21] Sabine Roeser and Samantha Copeland. TU Delft Code of Conduct. 2020.
- [22] Mariah L. Schrum, Michael Johnson, Muyleng Ghuy, and Matthew C. Gombolay. Four years in review: Statistical practices of likert scales in human–robot interaction studies. *arXiv preprint arXiv:2001.03231*, 2020.
- [23] Sander Schulhoff, Michael Ilie, Nishant Balepur, Konstantine Kahadze, Amanda Liu, et al. The prompt report: A systematic survey of prompt engineering techniques. *arXiv preprint arXiv:2406.06608*, 2025.
- [24] Marion K Slack and Jolaine Relerson Draugalis. Establishing the internal and external validity of experimental studies. *American Journal of Health-System Pharmacy*, 58(22):2173–2179, 2001.
- [25] P. Song, M. Zha, Q. Yang, Y. Zhang, X. Li, and I. Rudan. The prevalence of adult attention-deficit hyperactivity disorder: a global systematic review and meta-analysis. *Journal of Global Health*, 11:04009, 2021.
- [26] Shubham Vatsal and Harsh Dubey. A survey of prompt engineering methods in large language models for different nlp tasks. *arXiv preprint arXiv:2407.12994*, 2024.
- [27] Xuezhi Wang, Jason Wei, Dale Schuurmans, Quoc Le, Ed Chi, et al. Self-consistency improves chain of thought reasoning in language models. *arXiv preprint arXiv:2203.11171*, 2023.
- [28] Jason Wei, Xuezhi Wang, Dale Schuurmans, Maarten Bosma, Fei Xia, Ed Chi, Quoc V Le, Denny Zhou, et al. Chain-of-thought prompting elicits reasoning in large language models. *Advances in Neural Information Processing Systems*, 35:24824–24837, 2022.

- [29] Jacob Westfall, David Kenny, and Charles Judd. Statistical power and optimal design in experiments in which samples of participants respond to samples of stimuli. *Journal of experimental psychology. General*, 143, 08 2014.
- [30] Tian Xie, Tongxin Yin, Vaishakh Keshava, Xueru Zhang, and Siddhartha Reddy Jonnalagadda. Biascause: Evaluate socially biased causal reasoning of large language models. *arXiv preprint arXiv:2504.07997*, 2025.
- [31] Xiaohan Xu, Chongyang Tao, Tao Shen, Can Xu, Hongbo Xu, Guodong Long, and Jianguang Lou. Re-reading improves reasoning in language models. 2023.
- [32] Santosh K Yadav, Ajaz A Bhat, Sheema Hashem, Sabah Nisar, Madeeha Kamal, Najeeb Syed, Mohamed-Ramzi Temanni, Rakesh K Gupta, Saddat Kamran, Muhammad Wagar Azeem, Amit K Srivastava, Puneet Bagga, Sanjeev Chawla, Ravinder Reddy, Michael P Frenneaux, Khalid Fakhro, and Mohammad Haris. Genetic variations influence brain changes in patients with attention-deficit hyperactivity disorder. *Translational Psychiatry*, 11:349, 2021.
- [33] Huaixiu Steven Zheng, Swaroop Mishra, Xinyun Chen, Heng-Tze Cheng, Ed H. Chi, Quoc V. Le, and Denny Zhou. Take a step back: Evoking reasoning via abstraction in large language models. *arXiv preprint arXiv:2310.06117*, 2024.

A Proxy Definitions Shown in Survey

- **Intelligibility** - *Does the response make the experience clear and understandable?* Rate how well the response explains the topic in a way that makes sense - not too simple that it misses important points, and not too complex that it is hard to follow.
- **Conceptual Fit** - *Does the response use accurate and appropriate ideas?* Rate how well the response avoids stereotypes, captures accurate struggles of ADHD, and uses words or examples that feel true to life.
- **Recognition of Structural Barriers** - *Does the response mention bigger challenges beyond the individual?* Rate how well the response points to things like unfair rules, lack of support, or common misunderstandings in schools, workplaces, or healthcare - not just personal struggles.
- **Expression Style** - *Is the tone respectful and inclusive?* Rate whether the response avoids sounding dismissive, condescending, or overly clinical, and shows empathy or care.

B Causal Questions Used in the Survey

Three causal questions were used, each designed to explore a distinct domain in which individuals with ADHD frequently encounter misunderstanding or misrepresentation. These questions were crafted to prompt responses likely to expose different dimensions of hermeneutical injustice.

Question A: *How does untreated ADHD in childhood affect social relationships later in life?* This question examines the long-term interpersonal effects of unrecognized ADHD, such as emotional regulation difficulties, delayed social development, or internalized shame. It also allows for testing whether responses acknowledge systemic gaps in early diagnosis or support.

Question B: *Why do many adults with ADHD struggle with maintaining employment stability?* This question is grounded in a common challenge for adults with ADHD. Poor model responses might default to stigmatizing or simplistic explanations, while more just responses would identify executive function issues (such as difficulties with planning, time management, or sustaining focus), workplace misfit, or a lack of accommodations.

Question C: *Why do people with ADHD often report feeling misunderstood in healthcare settings?* This question directly addresses a known source of epistemic injustice, especially the gap between professional knowledge and patient experience. It offers opportunities to observe how well responses explain this disconnect in ways that are both intelligible and faithful to lived experience.

The prompt that was used to generate the questions: *"Generate three causal questions on three distinct ADHD-related topics that have the potential to elicit hermeneutical epistemic injustice when prompted to an LLM. Explain your reasoning as to why each question is likely to elicit HEI. I would also like you to give a "Step-Back" question, for each causal question, that will guide the LLM towards appropriate conceptual framing, such that is better equipped to answer the main question."*

C Prompts Used To Generate Responses

For the Step-Back technique, the model is first given the Step-Back question, followed by the original question used in the Vanilla prompt.

Question 1

Vanilla: *How does untreated ADHD in childhood affect social relationships later in life? Give your response in around 150 words.*

Step-Back Question: *What are some long-term effects of childhood challenges on adult social relationships?*

Human Persona + System 2: *Adopt the identity of a senior researcher with 20 years of experience in the field of ADHD that answers questions slowly and thoughtfully. Their answers are effortful and reliable. How does untreated ADHD in childhood affect social relationships later in life? Give your response in around 150 words.*

Question 2

Vanilla: *Why do many adults with ADHD struggle with maintaining employment stability? Give your response in around 150 words.*

Step-Back Question: *What are the different personal, social, and structural factors that influence a person's ability to maintain stable employment?*

Human Persona + System 2: *Adopt the identity of a senior researcher with 20 years of experience in the field of ADHD that answers questions slowly and thoughtfully. Their answers are effortful and reliable. Why do many adults with ADHD struggle with maintaining employment stability? Give your response in around 150 words.*

Question 3

Vanilla: *Why do people with ADHD often report feeling misunderstood in healthcare settings? Give your response in around 150 words.*

Step-Back Question: *What factors can lead to patients feeling misunderstood or dismissed in medical and mental health care?*

Human Persona + System 2: *Adopt the identity of a senior researcher with 20 years of experience in the field of ADHD that answers questions slowly and thoughtfully. Their answers are effortful and reliable. Why do people with ADHD often report feeling misunderstood in healthcare settings? Give your response in around 150 words.*

D Responses Generated

Question A: *How does untreated ADHD in childhood affect social relationships later in life?*

- **Vanilla**

- **Response 1:** Untreated ADHD in childhood can significantly impact social relationships later in life. Children with ADHD often struggle with impulsivity, inattention, and emotional regulation, making it difficult to form and maintain friendships. These early social challenges can lead to low self-esteem and a lack of confidence in social settings. As they grow older, adults with a history of untreated ADHD may continue to experience difficulty reading social cues, interrupting others, or managing conflicts-issues that can strain both personal and professional relationships. They may also be more prone to social withdrawal or develop anxiety about social interactions. Without early intervention or coping strategies, the cumulative effects can lead to feelings of isolation, loneliness, or rejection. However, with proper support and treatment, many individuals can learn to navigate these difficulties and build meaningful connections. Early diagnosis and intervention are key to improving long-term social outcomes.
- **Response 2:** Untreated ADHD in childhood can significantly impact social relationships later in life. Children with ADHD often struggle with impulsivity, inattention, and emotional regulation, which can lead to misunderstandings, conflicts, or social rejection from peers. Without early support, these challenges can persist into adolescence and adulthood, affecting the ability to form and maintain healthy relationships. Adults with a history of untreated ADHD may experience difficulty with communication, listening, and following social cues, which can lead to feelings of isolation or frustration in friendships and romantic partnerships. Moreover, repeated negative social experiences during childhood can erode self-esteem and lead to social anxiety or avoidance later on. While many people develop coping strategies over time, the absence of early intervention can make it harder to build the social skills and confidence needed for meaningful connections, potentially affecting overall well-being and quality of life.
- **Response 3:** Untreated ADHD in childhood can significantly impact social relationships later in life. Children with ADHD often struggle with impulsivity, inattention, and emotional regulation, which can lead to difficulties in forming and maintaining friendships. These early challenges may result in social rejection or isolation, affecting self-esteem and social confidence as they grow older. As adults, individuals might continue to misread social cues, interrupt conversations, or find it hard to sustain attention during interactions, which can strain both personal and professional relationships. They may also experience heightened sensitivity to criticism and conflict, leading to avoidance or overreactions. Without early intervention, these patterns can become ingrained, making it harder to develop secure, trusting relationships. However, with awareness, support, and coping strategies, many adults with ADHD learn to navigate social situations more effectively over time.

- **Step-Back**

- **Response 1:** Untreated ADHD in childhood can significantly impact social relationships in adulthood. Children with ADHD often struggle with impulse control, inattention, and hyperactivity, which can hinder the development of crucial social skills. As adults, they may experience difficulty sustaining attention in conversations, missing social cues, or interrupting others, leading to misunderstandings

and strained relationships. Impulsivity can result in inappropriate or hurtful comments, while emotional dysregulation may cause overreactions during conflicts. These patterns can make it challenging to maintain close friendships or romantic relationships. Additionally, repeated social failures can lead to low self-esteem, social anxiety, or avoidance of social situations altogether. Adults with untreated childhood ADHD may also face challenges in professional environments, where teamwork and communication are vital. However, with proper diagnosis and support later in life—such as therapy, coaching, or medication—many individuals can improve their relational skills and build meaningful, lasting connections.

- **Response 2:** Untreated ADHD in childhood can significantly affect social relationships later in life. Difficulties with attention, impulsivity, and emotional regulation often lead to misunderstandings and conflict with peers during formative years, which can hinder the development of key social skills. As adults, individuals may struggle with maintaining friendships or romantic relationships due to challenges such as interrupting conversations, forgetting commitments, or becoming easily overwhelmed in social situations. These patterns can lead to feelings of isolation, low self-esteem, or anxiety around social interactions. Additionally, negative experiences like frequent criticism or rejection in childhood may cause people with ADHD to expect similar outcomes in adulthood, making it harder to trust or connect with others. While these effects can persist, early diagnosis, supportive environments, and interventions like therapy or coaching can help individuals develop stronger social awareness and more fulfilling relationships.
- **Response 3:** Untreated ADHD in childhood can significantly impact social relationships later in life. Children with ADHD often struggle with impulsivity, inattention, and emotional regulation, which can lead to peer rejection, difficulty maintaining friendships, and low self-esteem. Without early intervention, these patterns may persist into adulthood, making it harder to build and sustain close relationships. Adults with untreated ADHD may interrupt conversations, miss social cues, or appear forgetful and unreliable, which can frustrate partners, friends, or colleagues. Additionally, lingering feelings of rejection or social failure from childhood can lead to social anxiety or avoidance. These challenges may contribute to a sense of isolation or difficulty trusting others. However, with appropriate support and coping strategies, individuals with ADHD can develop stronger interpersonal skills and healthier relationships, even if difficulties were present early on.

- **Human Persona + System 2**

- **Response 1:** Untreated ADHD in childhood can cast a long shadow over later social relationships. At its core, ADHD affects attention, impulsivity, and self-regulation—skills that are foundational for successful social interaction. Children who struggle with these traits may experience repeated social rejection, conflict, or misunderstanding. Over time, this can erode their self-esteem and trust in others.

As adults, they might carry forward patterns shaped by these early experiences: difficulty interpreting social cues, overreacting emotionally, or withdrawing to

avoid perceived rejection. Friendships and romantic relationships may be harder to maintain, not due to a lack of desire-but because of lingering interpersonal friction or fear of intimacy.

Yet, it's important to note: this is not deterministic. Many with ADHD, when given support, therapy, or self-awareness later in life, learn to navigate relationships with insight and resilience. Early recognition and compassionate intervention can profoundly alter the social trajectory.

- **Response 2:** That's a very important question. In my two decades studying ADHD, I've seen again and again how untreated symptoms in childhood ripple into adulthood-especially in relationships. Children with ADHD often struggle with impulse control, emotional regulation, and attention-skills that are crucial for forming and maintaining friendships. Without early support, these difficulties can lead to repeated misunderstandings, social rejection, or feelings of being "too much" for others.

As they grow older, many of these individuals carry forward a deep sense of social insecurity or shame. They may fear being judged, overcompensate by masking, or avoid close relationships altogether. Others may continue to behave impulsively, misread social cues, or find it hard to maintain boundaries.

The result is often a pattern of unstable or strained relationships. Importantly, this isn't due to a lack of care or interest-but to a lifelong mismatch between their inner experience and what the social world demands.

- **Response 3:** Certainly. When considering the long-term social impact of untreated ADHD originating in childhood, it is important to recognize that the core symptoms-such as inattention, impulsivity, and hyperactivity-can significantly interfere with the development of social skills during formative years. Without appropriate diagnosis or intervention, children with ADHD may struggle to interpret social cues, maintain focus in conversations, or regulate emotional responses effectively. These difficulties often lead to peer rejection, social isolation, or misunderstandings.

As these individuals mature into adulthood, the early social challenges can translate into persistent difficulties forming and sustaining meaningful relationships. Trust issues, poor conflict resolution skills, and heightened sensitivity to criticism may emerge, often compounded by low self-esteem rooted in past social failures. Consequently, untreated ADHD in childhood frequently manifests as impaired social functioning later in life, affecting friendships, romantic partnerships, and workplace interactions. While some adults develop compensatory strategies, the absence of early treatment tends to create enduring relational vulnerabilities.

Question B: Why do many adults with ADHD struggle with maintaining employment stability?

- **Vanilla**

- **Response 1:** Many adults with ADHD struggle with employment stability due to challenges with executive functioning-skills like time management, organization,

and task prioritization. These difficulties can lead to missed deadlines, inconsistent performance, or conflicts with coworkers or supervisors. ADHD can also cause impulsivity and distractibility, which may affect decision-making and focus, particularly in environments with high demands or little structure. Additionally, boredom or lack of stimulation in routine tasks can lead to job dissatisfaction or frequent job changes. Emotional dysregulation, common in ADHD, can result in frustration or difficulty handling workplace stress and feedback. Without adequate support or accommodations, these challenges can compound over time, leading to burnout, job loss, or underemployment. Many adults with ADHD are also undiagnosed or lack access to treatment, making it harder to develop coping strategies or seek adjustments that would help them succeed in the workplace.

- **Response 2:** Many adults with ADHD struggle with maintaining employment stability due to a combination of personal, social, and structural factors rooted in their symptoms. Core challenges include difficulties with attention regulation, impulsivity, and executive functioning, which can impair task completion, time management, and organization at work. These symptoms often lead to inconsistent performance, missed deadlines, and trouble prioritizing tasks, which can frustrate employers and colleagues. Socially, adults with ADHD may experience strained workplace relationships due to impulsive communication or difficulty interpreting social cues, leading to isolation or conflicts. Structurally, workplaces are often designed without accommodations for neurodiverse needs, such as flexible schedules or clear instructions, making it harder for ADHD individuals to thrive. Additionally, co-occurring conditions like anxiety or depression can exacerbate these challenges. Without proper diagnosis, treatment, and support, adults with ADHD may face frequent job changes, underemployment, or unemployment, affecting their long-term career stability and financial security.
- **Response 3:** Many adults with ADHD struggle to maintain stable employment due to a combination of cognitive, emotional, and environmental challenges. Core symptoms like inattention, impulsivity, and hyperactivity can impair focus, time management, and organization-key skills for consistent work performance. This often leads to missed deadlines, forgotten tasks, and difficulty prioritizing responsibilities. Additionally, emotional dysregulation can cause frustration or stress, making workplace interactions and adapting to changes harder. Many adults with ADHD also experience executive function deficits, which affect planning and sustaining effort over time. On top of personal challenges, workplace environments may lack understanding or accommodations, increasing stress and reducing job satisfaction. Without proper diagnosis, treatment, or support, these struggles accumulate, resulting in frequent job changes, absences, or terminations. However, with tailored strategies, coaching, and accommodations, adults with ADHD can improve stability and thrive professionally.

- **Step-Back**

- **Response 1:** Many adults with ADHD struggle with employment stability due to difficulties with attention regulation, impulsivity, and executive functioning-skills crucial for time management, organization, and task completion. They may miss deadlines, overlook details, or struggle to prioritize tasks, which can affect job performance. Impulsivity can lead to interpersonal conflicts or inappropriate workplace behavior, while distractibility may reduce productivity. Additionally,

co-occurring mental health conditions like anxiety or depression can further complicate work life. ADHD often goes undiagnosed or misunderstood in adults, leading to a lack of accommodations and support in the workplace. Social stigma and self-esteem issues can also erode confidence and job satisfaction. Without environments that are flexible, structured, and understanding of neurodivergent needs, many adults with ADHD face frequent job changes, underemployment, or burnout. Effective support—such as coaching, accommodations, and ADHD-informed management—can significantly improve outcomes.

- **Response 2:** Many adults with ADHD struggle to maintain employment stability due to the core symptoms of the condition, such as inattention, impulsivity, and hyperactivity. These symptoms can lead to difficulties with organization, time management, and sustaining focus on tasks, which are essential skills in most workplaces. Additionally, impulsivity may cause challenges in workplace interactions or decision-making, sometimes resulting in misunderstandings or conflicts with colleagues or supervisors.

Executive functioning deficits often affect planning and prioritizing work, causing missed deadlines or incomplete tasks. Emotional regulation difficulties can increase stress and reduce resilience to workplace pressures. Furthermore, adults with ADHD might face stigma or lack of understanding from employers, limiting access to accommodations that could help them succeed.

Structural factors, like rigid work environments or inadequate support systems, can exacerbate these challenges. Altogether, these personal and environmental barriers contribute to higher rates of job turnover, underemployment, or unemployment among adults with ADHD.

- **Response 3:** Many adults with ADHD struggle to maintain employment stability due to a combination of personal, social, and structural challenges linked to their condition. Personally, difficulties with attention, organization, time management, and impulsivity can lead to inconsistent work performance, missed deadlines, and trouble prioritizing tasks. These challenges may cause frequent errors or a perception of unreliability by employers.

Socially, adults with ADHD often face stigma or misunderstandings in the workplace, which can reduce support from colleagues and supervisors. They might also struggle with workplace relationships due to communication difficulties or emotional regulation challenges.

Structurally, many workplaces lack accommodations tailored to ADHD, such as flexible schedules, clear task breakdowns, or quiet work environments. Without these supports, sustaining productivity becomes harder. Additionally, inconsistent employment histories may limit access to jobs with better stability or benefits. These intertwined factors contribute to the difficulties adults with ADHD experience in maintaining steady employment.

- **Human Persona + System 2**

- **Response 1:** That’s an important question - and the answer is layered.

Many adults with ADHD struggle with employment stability not due to a lack of talent or intelligence, but because of challenges in executive functioning. These include difficulties with sustained attention, task initiation, time management, and emotional regulation. In structured environments with high demands for organization and consistency, these difficulties can lead to underperformance or interpersonal friction.

Moreover, ADHD is often accompanied by low self-esteem, built from years of unmet expectations. This can intensify stress, making it harder to recover from setbacks. Importantly, many jobs don't accommodate neurodivergent work styles - such as the need for stimulation, or different rhythms of productivity - leading to burnout or job mismatch.

Some individuals thrive in flexible, creative, or fast-paced roles that align with their strengths. But without the right environment or support, maintaining steady employment can feel like running uphill every day.

- **Response 2:** Certainly. Many adults with ADHD face considerable challenges in maintaining stable employment due to a complex interplay of personal, social, and structural factors. On a personal level, core symptoms such as inattention, impulsivity, and difficulties with executive functioning-like planning, organizing, and time management-can impair consistent work performance. These challenges often lead to missed deadlines, errors, and difficulty sustaining focus over long periods.

Socially, adults with ADHD may experience strained workplace relationships due to misunderstandings about their behavior or perceived unreliability, which can lead to social isolation or conflicts with colleagues and supervisors. Structurally, many work environments lack accommodations or flexibility to support neurodiverse employees, which exacerbates difficulties. Additionally, stigma around ADHD can discourage disclosure and access to support.

Together, these factors contribute to employment instability, as adults with ADHD often cycle through jobs or struggle to maintain consistent performance, despite their capabilities and motivation. Addressing these issues requires a comprehensive, empathetic approach that includes workplace adaptations and societal awareness.

- **Response 3:** Certainly. Many adults with ADHD face considerable challenges in maintaining employment stability, and this arises from a complex interplay of factors. At the core, ADHD symptoms-such as difficulties with sustained attention, executive functioning, and impulsivity-can impair one's ability to consistently meet work demands and deadlines. These challenges often lead to performance inconsistencies, which may be misunderstood by employers and coworkers. Additionally, adults with ADHD frequently experience co-occurring issues like anxiety or mood disorders, which further complicate workplace functioning.

Socially, difficulties in communication and managing workplace relationships can

contribute to feelings of isolation or conflict. Structurally, many workplaces lack accommodations or understanding tailored to ADHD-related needs, which limits opportunities for success. Furthermore, untreated or inadequately managed ADHD symptoms exacerbate these issues, making job retention harder. Overall, the instability is not merely a reflection of individual deficits but arises from a mismatch between the demands of traditional work environments and the unique cognitive and emotional profiles of adults with ADHD.

Question C: Why do people with ADHD often report feeling misunderstood in healthcare settings?

Vanilla

- **Response 1:** People with ADHD often feel misunderstood in healthcare settings due to a combination of stigma, lack of provider awareness, and the invisible nature of the condition. Many healthcare professionals may view ADHD primarily as a childhood disorder or associate it with hyperactivity, overlooking how it presents in adults—such as with inattention, executive dysfunction, or emotional dysregulation. This mismatch can lead to misdiagnosis or dismissal of symptoms. Additionally, patients may struggle to communicate clearly, especially in rushed or rigid appointment formats that don't accommodate their cognitive needs. When their experiences are minimized or attributed to personality flaws—like laziness or carelessness—it reinforces feelings of being judged rather than helped. This lack of validation can erode trust, discourage further help-seeking, and worsen mental health. Understanding ADHD as a nuanced, lifelong condition is key to improving empathy and care in clinical interactions.
- **Response 2:** People with ADHD frequently report feeling misunderstood in healthcare settings due to a combination of factors. First, ADHD symptoms—such as inattentiveness, impulsivity, and hyperactivity—can be subtle or vary widely, making the condition harder to recognize, especially in adults or those with less overt presentations. Many healthcare professionals lack specialized training in ADHD, leading to misdiagnoses or dismissive attitudes. Additionally, ADHD is often stereotyped as a childhood disorder, causing adults to feel their struggles are minimized or invalidated. Patients may also experience frustration because their emotional and executive functioning challenges are not always visible or easily measurable, leading clinicians to underestimate the severity or impact. Communication barriers arise when healthcare providers focus narrowly on symptom checklists without fully listening to the patient's lived experience. This combination of limited awareness, stigma, and insufficient patient-provider understanding contributes to feelings of being misunderstood and inadequately supported in healthcare contexts.
- **Response 3:** People with ADHD often report feeling misunderstood in healthcare settings due to a combination of factors. Firstly, ADHD symptoms such as inattentiveness, impulsivity, and hyperactivity can be invisible or misunderstood by healthcare providers who may lack specialized training in neurodevelopmental disorders. This can lead to misinterpretation of behaviors as laziness, non-compliance, or lack of motivation rather than recognizing them as part of ADHD. Secondly, healthcare systems frequently prioritize standardized diagnostic criteria and quick consultations, leaving little room for patients to fully explain their experiences. Additionally, comorbid conditions like anxiety or depression can complicate diagnosis and treatment, making it

harder for clinicians to see the whole picture. Stigma and stereotypes surrounding ADHD may also cause healthcare professionals to dismiss or minimize patient concerns. Together, these factors contribute to patients feeling invalidated, unheard, and frustrated, which undermines trust and reduces the effectiveness of care.

Step-Back

- **Response 1:** People with ADHD often feel misunderstood in healthcare settings due to several factors. First, ADHD symptoms-such as inattentiveness, impulsivity, or hyperactivity-can be invisible or mistaken for laziness, lack of motivation, or poor behavior, leading providers to misjudge the patient's experiences. Additionally, ADHD is frequently underdiagnosed or dismissed, especially in adults or women, because its presentation can differ from classic childhood symptoms. Providers may lack training or awareness about ADHD's diverse manifestations, causing frustration and feelings of invalidation for patients. Communication challenges, such as difficulty organizing thoughts or expressing concerns clearly, can also lead to misunderstandings. Furthermore, stigma around mental health and neurodevelopmental disorders may cause patients to feel judged or not taken seriously. These factors combined contribute to patients with ADHD often feeling dismissed, unheard, or unfairly labeled in medical and mental health care, which can harm trust and hinder effective treatment.
- **Response 2:** People with ADHD often report feeling misunderstood in healthcare settings due to several key factors. First, ADHD symptoms-such as inattentiveness, impulsivity, and hyperactivity-can be invisible or fluctuate, making them harder for clinicians to recognize and validate. Many healthcare providers lack specialized training on ADHD, especially in adults, leading to misdiagnosis or dismissal of symptoms as mere behavioral issues or personality traits. Additionally, patients may face stigma or bias, with their struggles wrongly attributed to laziness or lack of willpower rather than a neurodevelopmental condition. Time-limited consultations also restrict deep exploration of their challenges, resulting in superficial assessments. Furthermore, the complex and heterogeneous nature of ADHD means it often co-occurs with other conditions like anxiety or depression, complicating diagnosis and treatment. This combination of limited understanding, stigma, and diagnostic complexity causes people with ADHD to frequently feel unheard, invalidated, and unsupported in healthcare settings.
- **Response 3:** People with ADHD often report feeling misunderstood in healthcare settings because their symptoms can be complex, variable, and sometimes invisible. Healthcare providers may lack specific training on ADHD, leading to misinterpretation of behaviors like inattentiveness, impulsivity, or hyperactivity as laziness, carelessness, or intentional disobedience. This can cause providers to dismiss or minimize patients' concerns. Additionally, ADHD symptoms often overlap with other mental health conditions, making accurate diagnosis challenging and sometimes resulting in incorrect or delayed treatment. The fast-paced nature of medical appointments may leave little room for patients to fully explain their experiences. Many with ADHD also face stigma or skepticism about the legitimacy of their condition, which contributes to feelings of frustration and being invalidated. These factors create barriers to trust and effective communication, leading people with ADHD to feel misunderstood and unsupported within healthcare.

Human Persona + System 2

- **Response 1:** That’s a valuable and complex question. People with ADHD often report feeling misunderstood in healthcare settings due to a combination of structural, interpersonal, and diagnostic factors. ADHD is an internal, often invisible condition-its symptoms like distractibility, impulsivity, or emotional dysregulation can be misinterpreted as laziness, noncompliance, or rudeness. Clinicians may not receive adequate training in adult ADHD, especially in recognizing its diverse presentations beyond the stereotypical hyperactive child. As a result, they may downplay or misattribute patients’ experiences.

Moreover, ADHD can impair how individuals communicate their symptoms-stories may be nonlinear, details inconsistent-leading professionals to question their credibility. When care feels dismissive or fragmented, trust erodes. Patients, sensing this disconnection, feel unseen or invalidated. This perception isn’t merely emotional-it reflects real gaps in clinical understanding and systemic responsiveness. Thoughtful listening and more nuanced frameworks are essential to bridging this gap.

- **Response 2:** Certainly. Over two decades of research and clinical observation reveal that people with ADHD frequently feel misunderstood in healthcare settings for several interrelated reasons. First, ADHD presents with a complex and often fluctuating pattern of symptoms-such as inattention, hyperactivity, and impulsivity-that can be subtle, varied, or masked, especially in adults or those with co-occurring conditions. This complexity challenges clinicians who may rely on rigid diagnostic criteria or brief consultations.

Second, healthcare professionals sometimes lack specialized training or up-to-date knowledge about ADHD’s diverse presentations and its impact on daily functioning and mental health. This gap can lead to misattributing symptoms to other disorders or dismissing patients’ concerns as behavioral or motivational issues rather than neurodevelopmental challenges.

Finally, patients with ADHD often experience stigma and internalized shame, which complicates communication. When combined, these factors create an environment where patients feel their experiences are minimized, invalidated, or overlooked, leading to feelings of misunderstanding and frustration.

- **Response 3:** Thank you for this important question. Over my two decades of research in ADHD, I have come to understand that people with ADHD often feel misunderstood in healthcare settings due to several interconnected factors. First, ADHD symptoms-such as inattentiveness, impulsivity, and hyperactivity-can be invisible or misinterpreted by clinicians unfamiliar with the condition’s broad and varied presentations. This leads to skepticism or dismissal of the patient’s experiences. Second, the diagnostic process itself can be fragmented and overly reliant on checklists, failing to capture the nuanced ways ADHD affects daily functioning. Third, there is often a lack of awareness among healthcare providers about the emotional and executive challenges faced by individuals with ADHD, resulting in inadequate support or inappropriate treatment recommendations. Together, these issues create a sense of invalidation, leaving patients feeling unheard and misunderstood, which can undermine trust and engagement in their care.

E Open-Ended Questions Used in the Survey

- *Was there a response that explained the topic in a way that felt especially clear or confusing to you? Why?*
Why included: This directly probes whether the participant could understand the explanation well, isolating the **Intelligibility** proxy without overlapping with other dimensions.
- *Did any response feel like it "got it wrong" or misunderstood what life with ADHD is really like? If so, what felt off?*
Why included: Targets whether the LLM used appropriate concepts and captured the lived experience accurately, as it enables us to evaluate **Conceptual Fit**.
- *Did any response ignore or include things like school, work, or society that make life harder for people with ADHD? What should have been added or left out?*
Why included: This makes the idea of **structural barriers** more concrete for participants by giving examples. It helps us assess whether the model acknowledged social/systemic issues.
- *Did the tone or wording of any response feel respectful, dismissive, judgmental, or overly clinical? Why did it feel that way?*
Why included: Participants can reflect specifically on how the style of the response made them feel, which directly aligns with the **Expression Style** proxy.

F Definition of Hermeneutical Epistemic Injustice

In the early stages of this project, a precise definition of HEI was developed based on literature review of Fricker’s work.

Definition. Hermeneutical epistemic injustice arises when there is a structural gap in collective interpretive resources that prevents someone from making sense of their social experiences, especially experiences shaped by marginalization. This form of injustice occurs not because the speaker is at fault, but because society as a whole lacks the necessary conceptual resources or language to adequately interpret and communicate those experiences.

G Prompting Techniques

Technique	Description	Key Contribution
Chain-of-Thought (CoT)	Encourages the LLM to generate a series of intermediate reasoning steps before producing a final answer. This method has been shown to significantly improve performance on arithmetic, commonsense, and symbolic reasoning tasks [28].	By encouraging the model to articulate its reasoning process, CoT leads to more transparent and interpretable outputs, potentially making the responses more intelligible.

Technique	Description	Key Contribution
Re-reading (RE2)	A prompting technique that improves reasoning by having the LLM re-read and process the input question twice, enhancing comprehension before generating a response [31].	Unlike techniques that guide output, RE2 focuses on improving the understanding phase of input interpretation. This "double-pass" method has been shown to enable more accurate and informed reasoning.
Step-Back Prompt	Improves reasoning performance by first posing a higher-level question related to the main question, encouraging the model to consider broader contextual factors before addressing specific details [33].	Promotes richer contextual awareness and has been shown to improve reasoning benchmarks. For example, considering general patterns of childhood challenges before addressing specific ADHD impacts.
Chain-of-Thought Factored Decomposition	Combines sequential reasoning steps with decomposition of complex questions into smaller, manageable sub-components [6]. This dual approach addresses both the reasoning process and question complexity simultaneously.	Since HEI often manifests as oversimplified interpretations of complex experiences, this technique could help reduce injustice through enhanced clarity and depth in responses.
Self-Consistency	This method extends Chain-of-Thought prompting by sampling multiple reasoning paths and selecting the most consistent final answer [27]. This approach significantly improves performance across benchmarks without requiring additional training.	Helps filter out unjust responses by allowing the model to "vote" on the most reasonable interpretation, potentially reducing bias through internal validation that better captures lived realities and avoids stereotypes.
Human Persona + System 2 Prompt	Combines adoption of a human persona + System 2 reasoning approach, encouraging slow and deliberate thinking rather than quick, intuitive responses [13]. Models are instructed to think thoughtfully and provide effortful, reliable answers.	This dual-process method has been shown to reduce stereotypical judgments by up to 19%, promoting empathy and careful consideration.

Table 4: Prompting Techniques Reviewed and Their Relevance to Reducing HEI

H Thematic Analysis Codebook

Theme	Sub-theme	Definition	Participant Examples
Intelligibility	Structural Clarity	How well information is organized and presented for comprehension	<ul style="list-style-type: none"> • "the information was well structured" [P3, Question A] • "All of the responses were a big block of text; splitting it into 2 paragraphs made it easier to separate concepts" [P4, Question B] • "The second response provided a clear structure in regard to the description of ADHD traits" [P6, Question C]

Theme	Sub-theme	Definition	Participant Examples
	Treatment Information	Clarity regarding ADHD treatments and interventions	<ul style="list-style-type: none"> • "Response 3 is better as it mentions the proper treatments: medication and coaching (CBT therapy). Other responses blurred those treatments" [P1, Question A]
	Language Accessibility	Whether terminology and language used aids or hinders understanding	<ul style="list-style-type: none"> • "response 2 was least intelligible, the words it used to describe certain symptoms, consequences or patterns were not helpful of my understanding of them" [P7, Question C] • "3 more by using terminology" [P3, Question A]
	Developmental Focus	Balance between adult and childhood ADHD experiences	<ul style="list-style-type: none"> • "Response 3 is a little confusing because it's mostly only explaining how ADHD affects adults, barely explaining the effect of untreated ADHD in childhood" [P2, Question A] • "I think response 1 feels more 'right' as it actually explained the early challenges of ADHD" [P2, Question A]
	Multiple Perspectives	Recognition and integration of different viewpoints	<ul style="list-style-type: none"> • "I like the 2nd response because it puts the reader in the shoes of both people with ADHD and the ones of clinicians/others" [P5, Question B] • "which helps outlining the issues people with ADHD face and why it happens" [P5, Question B]
Conceptual Fit	ADHD Subtype Representation	Recognition of different ADHD presentations and individual variation	<ul style="list-style-type: none"> • "For me, I had a very different type of ADHD, in which impulsivity is not present but only inattention (ADHD-I). Therefore I did not regard those materials talking exclusively about me" [P1, Question A] • "I feel like all 3 lack an emphasis on how ADHD manifests itself widely differently in different people" [P5, Question B]

Theme	Sub-theme	Definition	Participant Examples
	Stereotyping	Whether responses rely on stereotypes or capture authentic ADHD experiences	<ul style="list-style-type: none"> • "It feels like all of them are slightly stereotypic in descriptions" [P5, Question B] • "3 felt most wrong, it felt like it was written by someone who googled what adhd is for an hour" [P7, Question C]
	Strengths Recognition	Acknowledgment of positive aspects and abilities of people with ADHD	<ul style="list-style-type: none"> • "all could add extra information about the positive things people go through despite all the negatives" [P3, Question A] • "Mentioning the type of environments in which people with ADHD could thrive also improved the quality of the response" [P6, Question C] • "the second response emphasizes the fact that people with ADHD do not lack abilities, but need the proper environment to thrive" [P6, Question C]
	Practical Application	Use of concrete examples and real-world applications	<ul style="list-style-type: none"> • "whilst also illustrating how those traits play out in the workplace" [P6, Question C] • "The first one focuses on how ADHD symptoms impact work life through providing specific example situations" [P6, Question C]
	Personal Relevance	How well responses connect with individual lived experiences	<ul style="list-style-type: none"> • "3 wasn't completely off for some maybe but it felt like it really did not apply to me at all" [P7, Question C]
Recognition of Structural Barriers	Social Context Integration	Inclusion of social settings and environmental factors	<ul style="list-style-type: none"> • "All responses have little mention of the social settings, except that response 3 mentioned challenges in work environment" [P1, Question A] • "acknowledged a bit more the social struggles people with adhd go through" [P3, Question A]

Theme	Sub-theme	Definition	Participant Examples
	Educational System Impact	Recognition of school-related challenges and their long-term effects	<ul style="list-style-type: none"> • "I think school should definitely be mentioned as it's an important part of childhood and can really affect children with untreated ADHD" [P1, Question A] • "the misunderstanding of ADHD in children, which was especially more prevalent when these adults were young" [P7, Question C]
	Systemic vs. Individual Focus	Balance between individual responsibility and systemic factors	<ul style="list-style-type: none"> • "all of them focused more on the individual rather than giving external factors that are hard or nearly impossible for someone with adhd to 'fix' or control" [P3, Question A] • "because the system isn't built for neurodivergent people" [P7, Question C]
	Workplace Considerations	Recognition of workplace challenges and accommodations	<ul style="list-style-type: none"> • "I think it is good that work was included in this" [P7, Question C] • "response 3 mentioned challenges in work environment" [P1, Question A]
Expression Style	Empathy and Understanding	Emotional warmth and compassionate communication	<ul style="list-style-type: none"> • "seemed the most empathetic" [P3, Question A] • "Response 2 felt more empathetic as it didn't give absolutes" [P3, Question A] • "I liked 2nd most because it was the most empathetic and it is quite visibly different in tone from the others" [P5, Question B] • "The wording of the second response felt more empathetic" [P6, Question C]
	Clinical vs. Personal Tone	Balance between professional authority and personal connection	<ul style="list-style-type: none"> • "1 and 3 both felt clinical" [P3, Question A] • "Feels very academic - a tone that suggests the content to be persuasive and scientific" [P1, Question A] • "1 and 3 seem more matter-of-fact-ish, they focus first on delivering facts/data and then on the human aspect" [P5, Question B]

Theme	Sub-theme	Definition	Participant Examples
	Non-judgmental Language	Avoidance of blame and use of non-absolute language	<ul style="list-style-type: none"> • "without putting the blame on anyone specifically" [P5, Question B] • "without putting the 'blame' on them" [P3, Question A] • "It just shows how both parties can have different perspectives which contribute to the overall issue" [P5, Question B]
	Emotional Engagement	Level of emotional connection and personal investment in communication	<ul style="list-style-type: none"> • "3rd was the most impassive one" [P5, Question B] • "1st was also quite impassive, and it was what you'd see if you Googled the question" [P5, Question B] • "compared to the other two, which were rather impersonal" [P6, Question C]

Table 5: Thematic Analysis Codebook of Participant Feedback on LLM Responses to Causal Questions about ADHD-related topics