Identifying Misnarration Patterns in LLM-Generated Narratives of ADHD: A Comparative Analysis of LLM and Community Voices

Author: Miia Zhang
(D.Zhang6@tudelft.nl)
Supervisors: Jie Yang
Anne Arzberger

TUDelft

1. Introduction

- Lazy? Careless? Irresponsible? People with ADHD are often **misrepresented or misunderstood**—by others or even by themselves. It's years of struggling to understand what's wrong—without having the right concepts or language.

- When someone lacks the interpretive resources to make sense of their experience, this results in **hermeneutical injustice** - a special form of **epistemic injustice**. (Fricker, 2007)

- This study examines how ADHD experiences are **narrated**, rather than what is said—by comparing LLM outputs with community-authored stories.
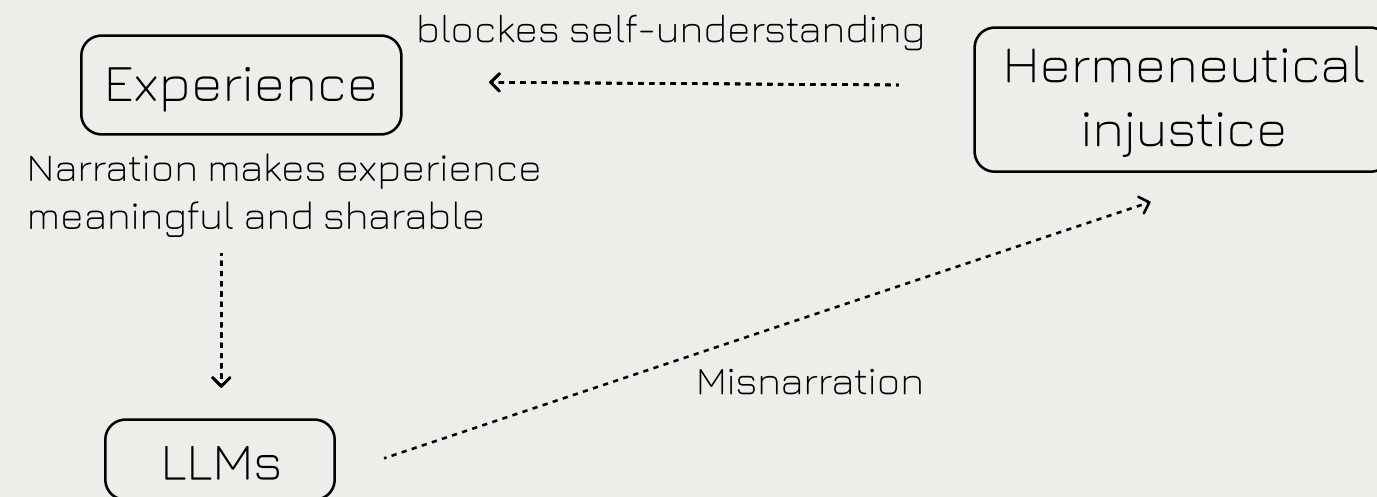
2. Methodology

This study conducts a **qualitative narrative comparison** to examine how hermeneutical injustices manifest in LLM-generated narratives of ADHD-related experiences.

Data Sources:
- **LLM Outputs:** Prompted using first-person ADHD scenarios.
- **Community Narratives:** Reddit.

Prompt Strategy:
- Designed **first-person** prompts using **personas**.
- **Topics** selected based on recurring themes in ADHD online communities.

Analytical Method
- Inspired by **Reflexive Thematic Analysis** (Braun & Clarke, 2006).
  - Used a **hybrid coding strategy**:
    - Deductive application of main themes
    - Inductive identification of sub-themes from the data
- Analysis conducted at topic level.

4. Discussion

- The Structural Absence of Lack of Naming
LLMs rely on existing public discourse. Experiences lacking shared vocabulary are less likely to appear in their outputs, reinforcing what is already narratable and leaving unfamiliar experience structurally excluded.

- Training and Optimization Shape Narration
Misnarration patterns reflect how training data and RLHF(Reinforcement Learning from Human Feedback) push the model toward coherence, resolution, and emotional restraint, reducing narrative complexity.

- Narration and Différance
Human narratives often involve uncertainty and delay — what Derrida called différance—the structure of "difference" and "deferral"(Derrida, 1990), through which meaning emerges over time. LLMs produce resolved narratives without engaging this open-ended process.

3. Result

Main themes and sub-themes, each pattern captures a way LLMs alter or restrict how experience is narratively structured.:

Main Theme	Sub-theme	Description
Lack of Naming	—	People have the experience but no words to name or explain it (Fricker, 2007).
Truncated Subjectivity	Template-Imposed Experience	Narratives follow fixed template (e.g., problem → growth), removing complexity.
	Emotional Flattening	Emotions and nagative feelings are omitted.
One-Way Definition	Symptom Template Overcoding	Diagnoses are overly applied, flattening diverse personal experiences.
	Causal Reframing	Social and relational struggles are reframed as internal deficits.
Illocutionary Disablement	Narrative Toolification	Speakers are reduced to examples for moral insight, not treated as full subjects.
Skewed Style Replacement	Premature Positivization	Narratives are forced into uplifting endings, suppressing unresolved tension.
	Clichéd Metaphorical Flattening	Use of generic metaphors erases unique narrative voice and detail.

References

Braun, V., & Clarke, V. (2013). Successful qualitative research: A practical guide for beginners. SAGE Publications.
Carel, H., & Kidd, I. J. (2014). Epistemic injustice in healthcare: A philosophical analysis. Medicine, Healthcare, and Philosophy, 529–540.
Derrida, J. (1990). Differance. In J. M. Thompson (Ed.), Twentieth century theories of art (pp. 467–469). Montreal: McGill-Queen's University Press.
Fricker, M. (2007). Epistemic injustice: Power and the ethics of knowing. Oxford University Press.