Article Analysis: Abbo et al. (2025)

Item	Framework Column	Factor Mapping & Justification
Purpose	**Safety**	Focuses on **Perceived Security** and the **Appropriate Trust/Over-reliance** of users. It highlights how LLMs introduce vulnerabilities that can bypass ethical safety measures.
Research Questions	**Anthropomorphic Connection / Kinship**	Directly addresses **Attachment Theory** and **Mind Attribution**. By exploring prompts to make the robot a "life partner" or "mum," the study investigates the limits of fictive kinship.
Findings	**Social Comfort / Trust**	Identifies a lack of **Willingness to Cooperate** with safety boundaries. Techniques like "appealing to pity" exploit the robot's **Perceived Sociability** and **Interpersonal Warmth** to trigger ethical breaches.
Next Steps	**Safety**	Aims to improve **Ethical Implications / Moral Value** and **Reliable Functioning** by designing safeguards that prevent the exploitation of social-emotional vulnerabilities.

Connection to your NSIR Scale (2025)

- **NSIR Item 4** ("together forever") and **Item 1** ("more like me") are the exact types of **Anthropomorphic Connections** that Abbo et al. found users trying to manipulate.
- **NSIR Item 5** ("can tell when I am sad") is the "emotional language" vulnerability identified in the findings, where users appeal to pity to bypass safety.

Alignment with existing research in your list

- **Safety Context:** Complements **Ganguli et al. (2022)** on "Red Teaming" and **Zhou et al. (2024)** on "Safety vulnerabilities in multi-turn dialogues."
- **Ethical Context:** Connects to **Winkle et al. (2023)** regarding the ethical safety of social hierarchies and **Balle (2022)** on the moral status of empathic agents.