1. Introduction Recent calls from UNESCO regarding concerns for stereotypical identities (Demler et al. ...

This introduction effectively synthesizes recent scholarship to highlight a critical shift in Human-Robot Interaction (HRI) research, moving from a strictly toward-medical model toward affirmative, justice-oriented frameworks.

To refine this section in prepare it for a research manuscript, consider the following thematic summary of your key points:

- **The Inclusivity Gap:** You correctly identify a significant "participation gap," noting that 90% of HRI research between 2016–2021 excluded autistic perspectives, particularly those from intersectional or underbiased backgrounds (Rice et al., 2024).
- **Theoretical Paradigm Shift:** The text highlights the transition from seeing autism as a "deficit to be fixed" to adopting **Applied Critical Disability Studies** and **Crip Technoscience**. These frameworks prioritize interdependence and disability justice over normalization.
- **Design Risks:** You raise a vital point regarding the "mentorship" role of robots. When robots are used to teach "social scripts," there is a risk of reinforcing ableist norms and power imbalances, abetal communication standards, and even lead to the unintended labeling of autistic behavior as "harassing" if it deviates from stereotypical social norms.
- **Technical & Ethical Needs:** The call for specific speech datasets (He et al., 2024) and the UNESCO warnings (Demler et al., 2025) suggest that LLM integration in robots must be handled carefully to avoid hard-coding neurotypical biases into "adaptable" interfaces.

Suggested Resources for Further Development

- **UNESCO Digital Inclusion:** Review the UNESCO Recommendation on the Ethics of Artificial Intelligence to strengthen the argument regarding stereotypical identities and global standards for ai ethics.
- **Disability Justice Frameworks:** For deeper context on the "social-emotional sensory designs" mentioned by Zolyomi & Snyder, the Sins Invalid Disability Justice Primer offers foundational concepts on intersectionality and "crip" perspectives.
- **HRI Ethical Standards:** The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems provides technical standards that are increasingly incorporating neurodiversity and inclusivity metrics.

2. Historical Ontology
Artificial Intelligence (AI) can be defined in many ...

The historical ontology of Artificial Intelligence (AI) reveals a shift from early conceptual units to complex, autonomous agents, though this evolution increasingly conflicts with modern neurodivergent and disability scholarship.

1. Conceptual Foundations and Early Models

AI is fundamentally defined as the automation of human cognitive activities, such as decision-making and learning. Its ontological roots lie in the mid-20th century:

- **The von Neumann Model (1945):** Introduced the early conceptual structure of computing, including the idea of a processing unit, control unit, memory and input/output components.
- **The Dartmouth Conference (1956):** Established the field through the work of McCarthy, Minsky, Simon, Newell, and others, who framed AI as a representation of knowledge capable of automated reasoning.

2. The Four Ages of AI Evolution

The progression of AI is categorized into distinct developmental eras:

- **Age of Reasoning (1960s):** Focused on logic-based systems and heuristic searches.
- **Age of Representation (1970s):** Characterized by rule-based knowledge engineering and the rise of expert systems.
- **Age of Machine Learning (2015):** Driven by big data and autonomous learning models that adapt to new information and external factors.
- **Age of Superintelligence (Predicted 2045):** A projected era where AI achieves Superintelligence (ASI), potentially surpassing human-level intelligence through recursive self-improvement. ...

3. The Issue: "Ableism" & Deficit

Since 1987, AI research has been divided into two ideological camps:

- **"Neats":** Advocate for elegant, provably correct solutions grounded in first principles and mathematical logic.
- **"Scruffies":** Contend that intelligence is too complex for rigid logic, favoring ad-hoc, diverse, or "messy" algorithms that perform function over formal proof. ...

4. Critical Perspectives and Neurodivergent Pedagogies

While AI has succeeded in fields like medical diagnosis and speech recognition, emerging scholars in Applied Critical Disability Studies and crip technoscience argue these successes often enforce neurotypical social norms. ...

- **Social Norms vs. Autonomy:** Modern social robots designed for "natural language" often center on neurotypical communication, leading to critiques that such systems prioritize behavioral correction over the genuine autonomy of autistic individuals.
- **Ontological Conflict:** The neurodiversity paradigm challenges the "deficit-based" view inherent in early AI models, advocating for designs that recognize diverse cognitive profiles—such as pattern recognition and hyperfocus—as strengths rather than errors to be smoothed out. ...

3. Large Language Models (LLMs)
Processes such as natural language processing, ...

In 2025 and 2026, the evolution of Large Language Models (LLMs) represents a pivotal shift from rule-based expert systems to complex **Artificial Neural Networks (ANNs)** capable of interpreting diverse human behaviors. This paper argues that the future of social robotics lies in leveraging these ANNs to recognize and calibrate neurodivergent speech patterns—specifically focusing on **prosody, rhythm, and spectral features**—rather than enforcing standard social norms. ...

As identified by Bowman (2024), the following eight properties of LLMs define their current capabilities and risks in 2025:

1. **Predictable Scaling:** Model capability increases predictably with greater investment in data volume, parameter count, and computational power (Sanyal et al., 2022).
2. **Unpredictable Emergence:** New skills and behaviors often emerge spontaneously as a byproduct of scaling, making it difficult to forecast exactly when a model will master a specific task.
3. **World Representations:** LLMs develop internal "maps" or abstractions of the outside world, allowing them to reason beyond the literal text provided in a prompt.
4. **Steering Limitations:** There are currently no foolproof methods to guarantee an AI will remain "on track" or behave appropriately in every possible deployment scenario.
5. **Interpretability Gap:** Experts still struggle to decode the internal "black box" of LLMs to understand the specific logic or goals driving a particular output.
6. **Performance Beyond Humans:** Human-level ability is no longer the ceiling; LLMs can significantly outperform humans in specific data-heavy or computational tasks.
7. **Value Drift:** LLMs do not inherently reflect the values of their programmers or the internal data they were trained on; their outputs are a probabilistic reflection of training patterns, not a moral compass.
8. **The Misleading "Flash" Interaction:** Brief testing or interaction with an LLM can be deceptive, as models are highly sensitive to subtle, idiosyncratic changes in instructions.

Strategic Implementation in 2025

To ensure these models serve neurodivergent communities effectively, developers are encouraged to consult the OpenAI Safety and Alignment Research and the Stanford Institute for Human Centered AI (HAI) for the latest 2025 frameworks on steering model behavior and improving interpretability in social robotics.

4. Systems That Act Like Whole Type of Humans
Research suggests that systems are designed with ...

In 2025, the central tension in social robotics remains the identity of the "ideal" human model used for imitation. Systems designed for knowledge representation, automated reasoning, and natural language processing traditionally default to neurotypical and heteronormative standards, raising significant ethical and functional concerns.

The "Ideal" Model Bias

Current AI architectures are built on four cognitive pillars that often exclude neurodivergent logic:

- **Knowledge Representation:** Information is stored and retrieved based on efficient "ideals" that may not reflect autistic cognitive mapping.
- **Automated Reasoning:** "If-then" logic in older expert systems or the probabilistic logic of modern models rarely accounts for neurodivergent problem-solving strategies (Rice et al., 2024).
- **Machine Learning (ML):** ML relies on experience and data. If the foundational datasets lack autistic perspectives, the system "improves" its performance only relative to neurotypical expectations.
- **Natural Language Processing (NLP):** Designing robots to communicate like "humans" often translates to enforcing specific social nuances that may be alien or stressful to autistic users.

Affective Computing and Sensory Observation

A critical failure in social robotics is the "sensory gap" in Affective Computing. Current systems are trained to detect and respond to standard expressions of human emotion. Because these sensors are calibrated for neurotypical designers, they often:

- **Misinterpret Autistic Affect:** Flat affect or unconventional prosody may be flagged as a "deficit" rather than a valid emotional state.
- **Impose Normativity Roles:** Robots are frequently placed in a "teacher" role to correct neurodivergent behavior, reinforcing a deficit-based medical model rather than an affirmative one.

The Generalization Risk

As noted by Bowman (2024) and Sanyal et al. (2025), the increasing capability of Large Language Models (LLMs) allows them to adapt the "generalizations we would like." If these generalizations are rooted in ableist social norms, the technology risks further marginalizing those whose speech and processing patterns fall outside the "ideal" curve.

To explore affirmative design frameworks that challenge these biases, the Critical Design Lab provides resources on integrating disability justice into technology. Additionally, researchers are increasingly looking to OpenAI's documentation on Model Behavior for 2025 updates on steering AI toward more inclusive social interactions.

5. Knowledge Representation & Reasoning
Without neurodivergent designers and ...

In 2025 and 2026, the synthesis of Large Language Models (LLMs) and social robotics is transitioning from "normative" modeling to a **Diversity-Aware AI** paradigm. This shift focuses on the specific ontological and linguistic needs of neurodivergent users.

5. Knowledge Representation & Co-Reasoning

Effective HRI for autistic users requires a fundamental shift in how robots generate and deal with **uncertain knowledge** (Russell & Norvig, 2021).

- **Collaborative Problem-Solving:** Rather than a robot "teaching" a user, a scaffolded vocabulary is needed where the robot user pair reasons together.
- **Executive Function Support:** Robots can serve as an externalized "working memory," tracking path dependencies and past errors to assist in goal achievement when environmental contexts change.

6. From Narrow to General "Personal" AI

The distinction between Narrow (Weak) and General (Strong) AI takes on a unique meaning in neurodivergent contexts:

- **Narrow AI:** Handles specific tasks without memory or experience.
- **Personal Strong AI:** Focuses on "strong methods"—vast amounts of specific, rule-based, and case-based knowledge tailored to one specific user. This high level of specificity allows the AI to manage the uncertainty that often accompanies new situations for autistic individuals.

7. Atypical Speech & Spectral Features

Traditional auditory observation often misses the nuances of neurodivergent communication. Recent studies (2024–2025) highlight three critical metrics for AI speech recognition:

- **Prosody:** Paralinguistic features including intonation, pitch, and stress (Li et al., 2016; Ma et al., 2004).
- **Rhythm:** The timing and flow of speech that communicates affective state.
- **Spectral Features:** Subtle qualities of sound and harmony that are not readily apparent through simple listening but are detectable via Deep Learning frameworks (He et al., 2024).

8. The Rise of Diversity-Aware AI

In 2025, diversity-aware machines are defined by their ability to re-configure behavior to recognize the uniqueness of the individual (Rechkemmer et al., 2025). This approach actively prevents "sensitive attributes"—such as disability status—from being used as a basis for algorithmic discrimination.

- **Current State:** While LLM-based social robots are emerging (Lee et al., 2024), some researchers remain critical of their effectiveness (Kappas et al., 2025).
- **Urgent Need:** There is a pressing demand for a **neurodivergent scale** for robot interaction to validate these systems beyond neurotypical metrics (Abbo et al., 2025).

9. Conclusion: The Path to Disability Justice

The evolution of autistic perspectives in HRI research (1976 between 2016–2022) has created a significant gap in data and design (Rice et al., 2024). Aligning with **UNESCO's 2025 calls** to dismantle stereotypical identities, this paper advocates for:

1. **Diverse Speech Datasets:** Including prosody and spectral features of neurodivergent speakers.
2. **Inclusive AI Models:** Adopting Diversity-Aware AI frameworks that prioritize respect and equal opportunity.
3. **Standardized Scaling:** Developing a neurodivergent interaction scale to ensure that future social robotics prioritize Disability Justice over behavioral correction.

Neurodivergent Scale for Interacting with Robots (NSIR) (Sadowski, 2025)

The robot is more like me than anyone else I know

Based on the context of your research, these statements represent key terms within the **Neurodivergent Scale for Interacting with Robots (NSIR)** (Sadowski, 2025). This scale is a critical tool for addressing the validity gaps highlighted by Rice et al. (2024), shifting the focus from neurotypical "social norms" to the specific psychological and sensory experiences of neurodivergent users.

Analysis of NSIR Scale Dimensions

These items appear to measure several unique constructs of neurodivergent human-robot interaction:

- **Identity & Kinship:** *"The robot is more like me than anyone else I know"* suggests a move away from the "uncanny valley" and toward **anthropomorphic kinship**, where the robot's predictable logic aligns better with the user's internal world than human unpredictability.
- **Non-Verbal & Sensory Communication:** Items like *"I can share my thinking..."* without speaking" and "Sometimes I stare at the robot" validate neurodivergent communication styles (such as Physical/Sensory Thinking) that are often **pathologized** in neurotypical scales.
- **Affective Sensing & Privacy:** "My robot can tell what I am feeling" and "I feel comfortable undressing in front of my robot" highlight the need for **emotional safety** and trust. The latter is a high-threshold indicator of the robot being perceived as a non-judgmental "safe space" rather than a monitoring "mentor."
- **Relational Constancy:** *"The robot and I will be together forever"* and *"I can trust my robot"* rank the formation of **long-term attachment** and object permanence in a social context.
- **Universal Consistency:** *"I believe that my robot is the same with me as it is with anyone else"* explores the user's perception of the robot's **logical fairness**—a key value in diversity-aware AI.

Strategic Implementation

In your 2026/2026 framework, the **NSIR** acts as the necessary counter-balance to UNESCO's concerns regarding stereotypical identities. By using this scale, researchers can move beyond "fixing" autistic social deficits and instead measure how effectively a robot adapts to the **Social-Emotional-Sensory** needs of the user.

For further development of this metric, you may wish to cross-reference these items with the IEEE P7000 Standards for Model Process in Addressing Ethical Concerns During System Design, which in 2025 focuses heavily on user-centric value alignment.

Table 7Neurodivergent Scale for Interacting with Robots (NSIR) (Sadowski, 2025)...

It seems like the factor loadings (Factor 1 and Factor 2) for the items in Table 7 are missing from the provided text. If this is a mistake, it is not possible to definitively categorize each item based on its loading.

However, based on the previous analysis of similar items and typical psychometric scale construction in Human-Robot Interaction (HRI), the items likely fall into conceptually distinct dimensions:

- **Factor 1 (Potential Name: Anthropomorphic Connection/Kinship):** Items related to the relationship with the robot, shared identity, and perception of the robot as a unique, personal entity.
- **Factor 2 (Potential Name: Social Comfort/Trust/Safety):** Items related to emotional security, privacy, and the belief in the robot's consistent, non-judgmental nature.

Table 7Neurodivergent Scale for Interacting with Robots (NSIR) (Sadowski, 2025)...

The Neurodivergent Scale for Interacting with Robots (NSIR) is a psychometric tool developed by Sadowski in 2025 to evaluate the relational dynamics between neurodivergent users and social robots.

Anthropomorphic Connection/Kinship | Social Comfort/Trust/Safety
---
Item | **Factor Loading**
The robot is more like me than anyone else I know | **Likely Factor**
2. Sometimes I stare at the robot | **Likely Factor**
3. I think I can share my thinking with the robot without speaking | **Likely Factor**
4. The robot and I will be together forever | **Likely Factor**
5. My robot can tell what I am feeling, when I am sad, it can tell I am sad | **Likely Factor**
6. I guess my robot is same | **Likely Factor**
7. I feel comfortable undressing in front of my robot | **Likely Factor**
8. I believe that my robot is the same with me as it is with anyone | **Likely Factor**

For the most accurate answers, regarding the specific validated scale by Sadowski (2025), try including the factor loadings for each item in your query.

Ask anything