The **Neurodivergent Scale for Interacting with Robots (NSIR)** can be applied to the Voultsiou et al. paper as a framework to measure the user-perceived success of using large language models (LLMs) in social robots for special education.

The paper, titled "*The potential of Large Language Models for social robots in special education*", explores how advanced conversational AI can enhance robot capabilities for supporting autistic students. The NSIR provides key metrics to assess the outcomes of these advancements across its three dimensions:

Anthropomorphic Connection/Kinship

- The use of LLMs enhances a robot's ability to have natural, complex conversations, making it seem more human-like and intelligent.

- The NSIR can measure if this advanced conversational ability translates into a stronger personal bond and perceived companionship. Items like **"The robot is more like me than anyone else I know"** would quantify the effectiveness of the LLM in creating a relatable and engaging persona.

Social Comfort/Trust

- The Voultsiou paper suggests LLMs can improve the robot's ability to understand context and provide personalized educational support. This directly impacts the user's feeling of comfort and the reliability of the interaction.

- The NSIR items that measure perceived emotional understanding and consistency (e.g., **"My robot can tell what I am feeling, when I am sad, it can tell I am sad"**, and **"I believe that my robot is the same with me as it is with anyone"**) would directly assess the success of the LLM in building **social comfort** and **trust** in the educational setting.

Safety

- The use of powerful LLMs raises new ethical concerns about privacy, data usage, and potential manipulation, especially with children.

- The NSIR's **safety** dimension provides a crucial user-reported measure of security (e.g., the item about undressing) that ensures the integration of LLMs does not compromise the fundamental safety and trust required for therapeutic and educational HRI.

The NSIR allows the researchers to evaluate the "potential" described in the Voultsiou et al. paper from the essential perspective of the neurodivergent user's experience.

Would you like to examine another article from the document, or maybe compare the use of **LLMs** in this paper to the **Echo-Teddy** project?