Before Words, Beyond Speech: Evaluating Nonverbal Social Reasoning in Early Childhood

Marie A. Huynh*, Laura Bravo-Sánchez*

Lauren Klein, Nick Haber, Philip Fisher, Serena Yeung-Levy

Stanford University

NEURAL INFORMATION PROCESSING SYSTEMS

Real world systems must understand how people respond to each other.

Social Understanding



Did the caregiver follow the child's lead?

Team coordination



Did the surgical team respond to a request for help?

Human-AI interaction



Should the robot help or wait?

Video links visible cues to responses across people and time.

Photos: Unsplash

Huynh, M.A.*, Bravo-Sánchez, L.*, Klein, L., Haber, N., Fisher, P., Yeung-Levy, S. Before Words, Beyond Speech: Evaluating Nonverbal Social Reasoning in Early Childhood. NeurIPS 2026

VLM → Social interaction reasoning



Q: “What happens between person red and green?”

A: “Person red and green are in a library. Person red points to person green’s book, person green frowns and pushes red. The adult comes over to intervene.”

Huynh, M.A.*, Bravo-Sánchez, L.*, Klein, L., Haber, N., Fisher, P., Yeung-Levy, S. Before Words, Beyond Speech: Evaluating Nonverbal Social Reasoning in Early Childhood. NeurIPS 2026

VLM → Social interaction reasoning



Q: “What happens between person red and green?”

A: “Person red and green are in a library. Person red points to person green’s book, person green frowns and pushes red. The adult comes over to intervene.”

Current evaluations hide failure modes.

Sim. Score: 1.5/3.0

A1: … Person green is reading a book while person red watches.

✗ wrong action

A2: … Person red points to person green’s book and person green smiles.

✗ wrong facial expression

Huynh, M.A.*, Bravo-Sánchez, L.*, Klein, L., Haber, N., Fisher, P., Yeung-Levy, S. Before Words, Beyond Speech: Evaluating Nonverbal Social Reasoning in Early Childhood. NeurIPS 2026

VLM → Social interaction reasoning



Q: “What happens between person red and green?”

A: “Person red and green are in a library. Person red points to person green’s book, person green smiles and pushes red. The adult comes over to intervene.”

❌ wrong facial expression

NEST benchmark

Social reasoning

- Open descriptions — Sim.: 1.5/3.0
- Interpersonal — I: 20%
- Behavioral — B: 50%
- Contextual — C: 80%

Huynh, M.A.*, Bravo-Sánchez, L.*, Klein, L., Haber, N., Fisher, P., Yeung-Levy, S. Before Words, Beyond Speech: Evaluating Nonverbal Social Reasoning in Early Childhood. NeurIPS 2026

Social interaction reasoning

1.  NEST benchmark

2. Thorough analysis

3. Real-world relevance

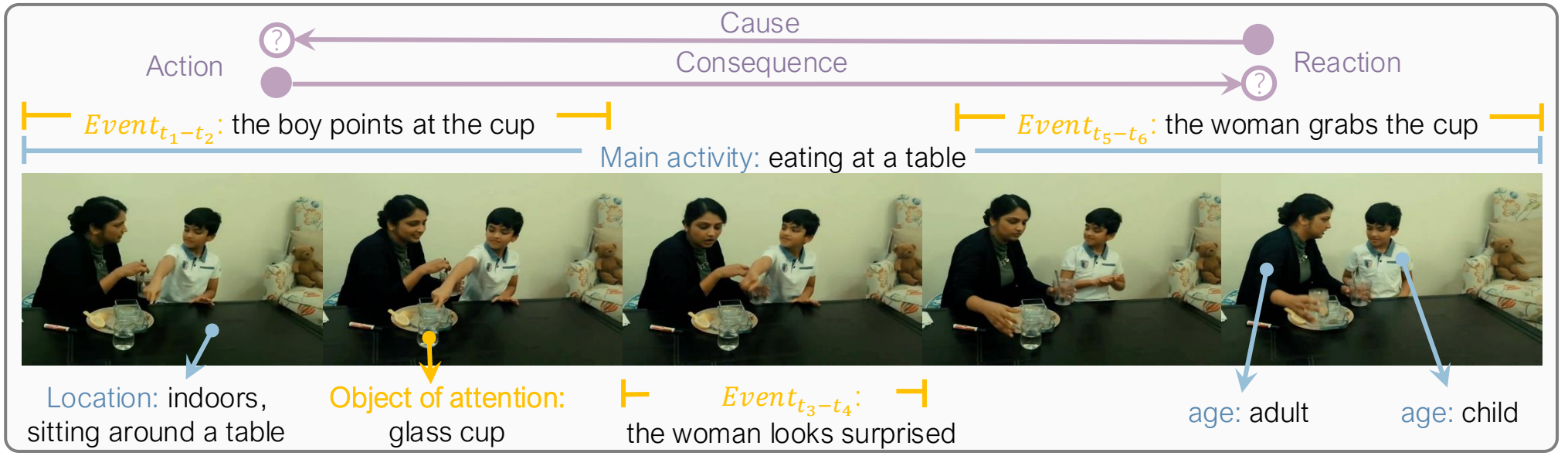
Huynh, M.A.*, Bravo-Sánchez, L.*, Klein, L., Haber, N., Fisher, P., Yeung-Levy, S. Before Words, Beyond Speech: Evaluating Nonverbal Social Reasoning in Early Childhood. NeurIPS 2026

1

NEST benchmark

How to structure non-verbal social interactions?

From one interaction …



… to 12 questions

Contextual			Behavioral			Interpersonal		Open interaction
age	location	main activity	shared attention	object of attention	event at time t	cause	consequence	description of the interaction

Huynh, M.A.*, Bravo-Sánchez, L.*, Klein, L., Haber, N., Fisher, P., Yeung-Levy, S. Before Words, Beyond Speech: Evaluating Nonverbal Social Reasoning in Early Childhood. NeurIPS 2026

1

NEST benchmark

Understanding children's social interactions beyond words

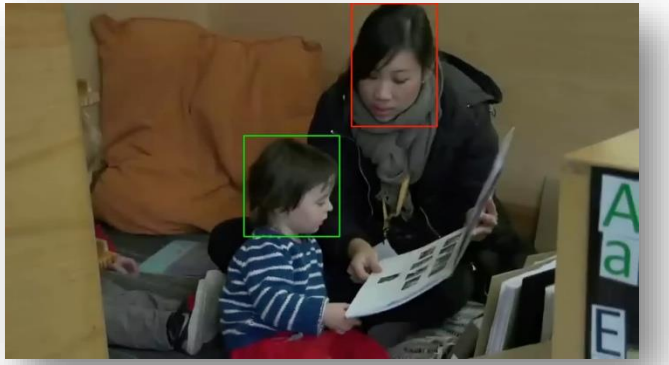
1,208 annotated clips

14,024 structured questions

10+ countries

150+ participants

Naturalistic



Non-verbal



Diverse



Huynh, M.A.*, Bravo-Sánchez, L.*, Klein, L., Haber, N., Fisher, P., Yeung-Levy, S. Before Words, Beyond Speech: Evaluating Nonverbal Social Reasoning in Early Childhood. NeurIPS 2026

Social interaction reasoning

1



NEST benchmark

2

Thorough analysis

3

Real-world relevance

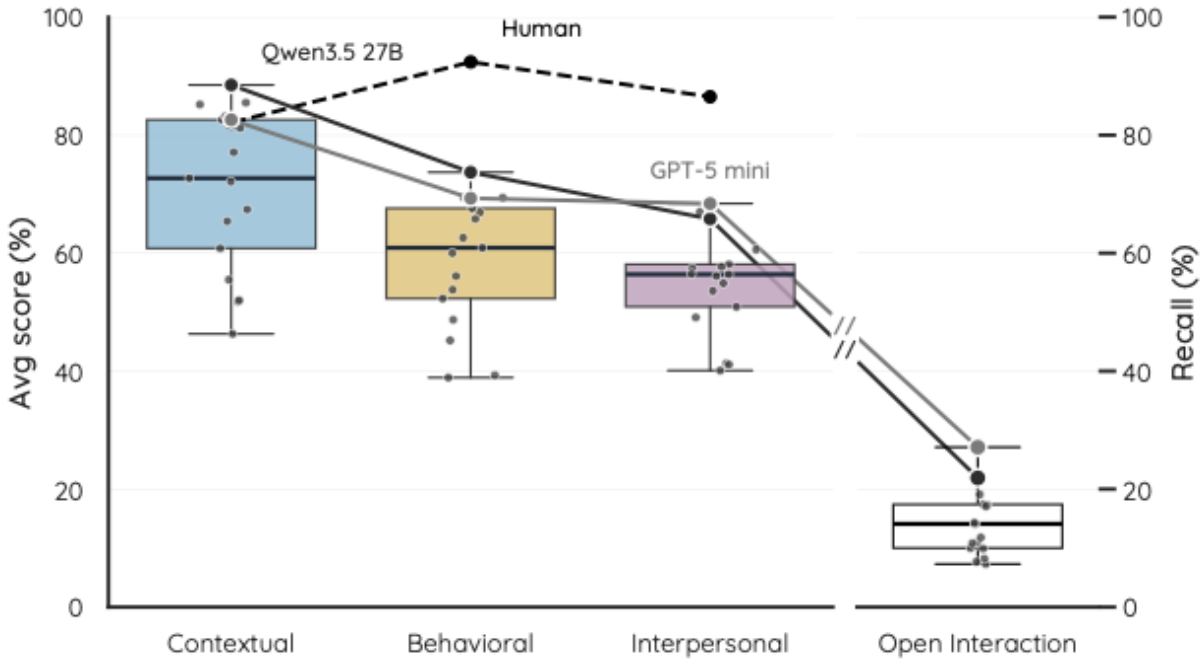
Huynh, M.A.*, Bravo-Sánchez, L.*, Klein, L., Haber, N., Fisher, P., Yeung-Levy, S. Before Words, Beyond Speech: Evaluating Nonverbal Social Reasoning in Early Childhood. NeurIPS 2026

2

Thorough analysis

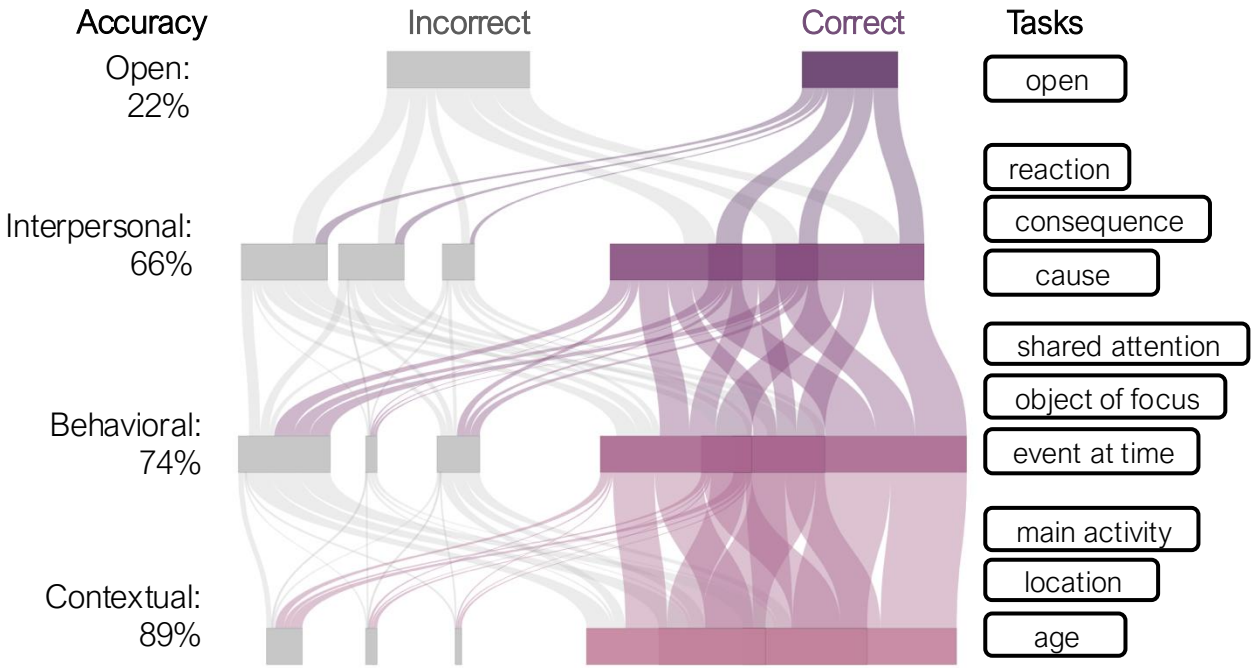
NEST

Compare across models



Models struggle with higher level social concepts

Delve into a model



Sankey plot of the best open-source model Qwen 3.5-27B

Huynh, M.A.*, Bravo-Sánchez, L.*, Klein, L., Haber, N., Fisher, P., Yeung-Levy, S. Before Words, Beyond Speech: Evaluating Nonverbal Social Reasoning in Early Childhood. NeurIPS 2026

Social interaction reasoning

1.  NEST benchmark

2. Thorough analysis

3. Real-world relevance

Huynh, M.A.*, Bravo-Sánchez, L.*, Klein, L., Haber, N., Fisher, P., Yeung-Levy, S. Before Words, Beyond Speech: Evaluating Nonverbal Social Reasoning in Early Childhood. NeurIPS 2026

3

Real-world relevance

Can models be used for social interaction reasoning in the wild?

NEST

Finetuning	Does domain specific data improve performance?
Specialized models	How do specialist models compare to generalist VLMs?
Case study	Do the findings transfer to a parent-coaching intervention?

FIND
Stanford Filming Interactions to Nurture Development

Huynh, M.A.*, Bravo-Sánchez, L.*, Klein, L., Haber, N., Fisher, P., Yeung-Levy, S. Before Words, Beyond Speech: Evaluating Nonverbal Social Reasoning in Early Childhood. NeurIPS 2026

Come find us in Paris and Sydney!

Scan for the project page

Links

Project page https://laubravo.github.io/nest_website/

Code https://github.com/yeung-lab/NEST

Benchmark https://huggingface.co/datasets/mahuynh/NEST

Find us at Sydney and Paris

Poster session 3 (Sydney)

Wed Dec 9th · 10am-1pm AEDT · Hall 1-4

Laura and **Marie** will be there in person in Paris and Sydney, respectively. Come find us during the conference!

With thanks to

FIND
Stanford Filming Interactions to Nurture Development

Stanford University