



OmniSapiens: A Foundation Model for Social Behavior Processing via Heterogeneity-Aware Relative Policy Optimization

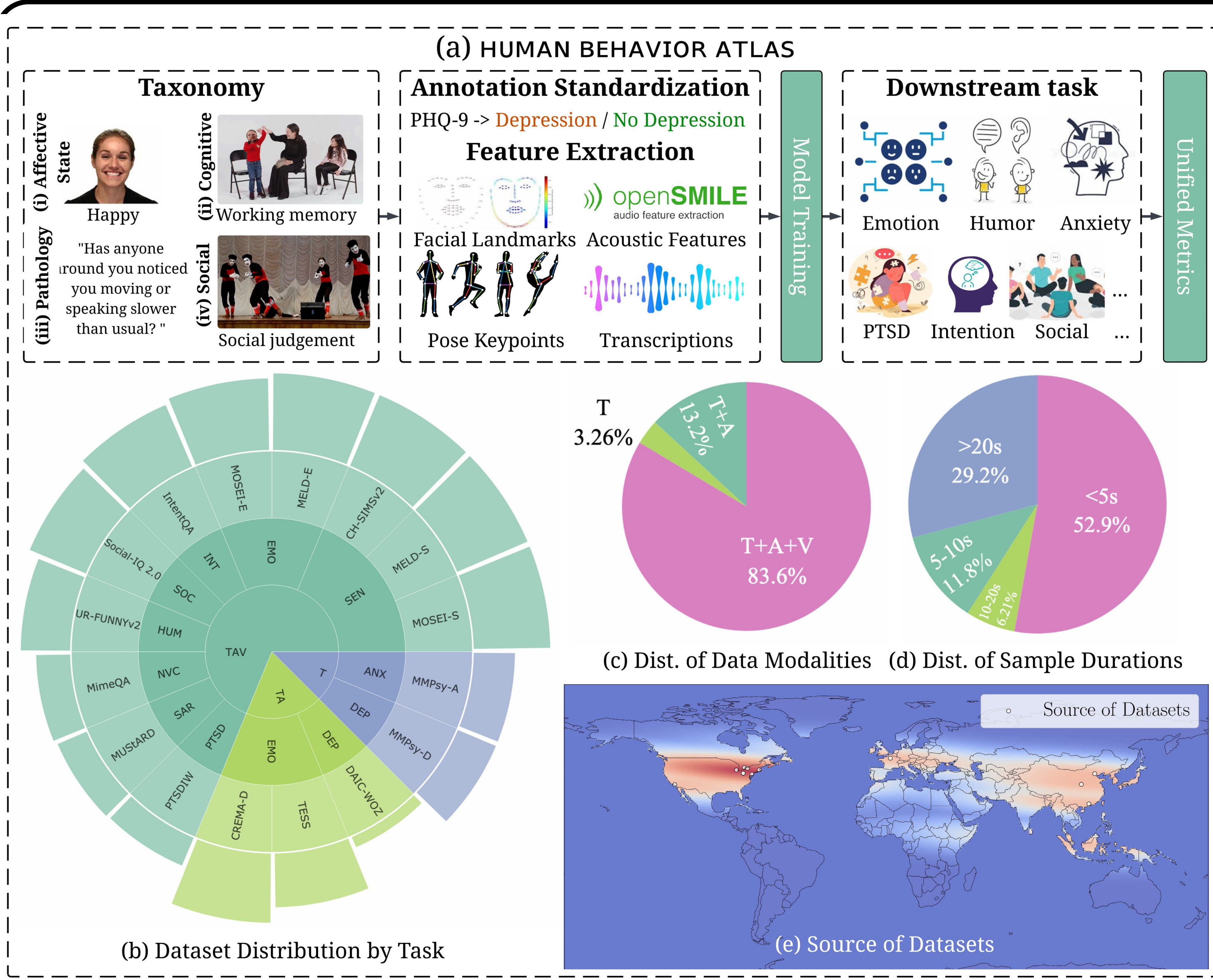
Keane Ong^{1,2}, Sabri Boughorbel³, Luwei Xiao¹, Chanakya Ekbote², Wei Dai², Ao Qu², Jingyao Wu², Rui Mao⁴, Ehsan Hoque⁵, Erik Cambria⁴, Gianmarco Mengaldo^{†1}, Paul Pu Liang^{†2}

¹National University of Singapore, Singapore, ²Massachusetts Institute of Technology, USA, ³Prince Sattam bin Abdulaziz University, ⁴Nanyang Technological University, Singapore, ⁵University of Rochester

Links to all
paper &
contact:



Human Behavior Atlas Benchmark



Building the Next Frontier of Socially Intelligent AI:

Social Intelligence does not emerge from an isolated understanding of behavior cues. Current AI → fragmented learning, poor transfer,

Our Goal: Build unified foundation models for human behavior that can reason and generalize across tasks and domains.

Overview: 10 Human Behavior Tasks; 101K+ samples; instruction sets compatible with foundation model training.

Outcome: OmniSapiens model family - first of its kind for Human Behavior Analysis!

RQ1 & RQ2 : Can Behavioral Foundation Models Perform Over Multiple Behavioral Tasks, and Generalize to Unseen Behavioral Contexts?

Models	EMO	HUM	INT	PTSD	ANX	DEP	SEN	SAR	SOC	NVC	Avg. Rank ↓
Gemma-3-4B	55.03	59.70	22.70	49.90	60.10	46.25	73.83	52.90	19.10	2.30	7.70
Qwen 2.5-Omni-7B	58.25	54.30	25.40	76.00	79.30	71.35	67.20	65.60	25.40	6.90	5.80
Qwen 2.5-VL-7B	54.08	58.30	24.90	75.50	63.10	63.80	50.50	51.10	23.10	9.80	7.30
Qwen 3-VL-8B-Instruct	57.66	66.76	38.00	92.70	42.29	51.62	69.70	63.67	24.94	13.95	5.30
HumanOmniV2-7B*	59.70	63.80	26.30	82.40	52.70	65.40	74.20	39.50	28.20	9.30	5.50
OmniSapiens SFT*	63.08	53.20	25.60	100.00	90.90	73.25	76.77	62.40	25.70	12.10	4.20
OmniSapiens BAM*	64.53	64.40	17.70	100.00	90.90	78.85	78.53	79.50	20.10	16.20	3.10
OmniSapiens-7B RL*	57.28	63.90	48.60	96.80	91.90	77.15	39.60	64.70	30.40	13.30	4.00
OmniSapiens-7B 2.0 (ours)	76.55	69.85	50.52	98.39	91.98	78.87	77.61	70.64	25.40	14.54	1.80

Model	AUT	SER	IDR	SMSA	SIR
Qwen 2.5-Omni-7B	25.68	53.53	70.25	44.64	34.99
Gemma 4-E4B-Instruct (8B)	16.78	58.30	11.37	55.60	56.74
OmniSapiens-7B RL	30.46	55.77	69.29	55.03	66.53
HumanOmniV2-7B	38.05	62.74	21.97	53.06	37.45
OmniSapiens-7B 2.0 (ours)	39.91	72.11	72.43	58.47	69.27

- **OmniSapiens 2.0** demonstrates performance gains by resolving learning imbalances
- **OmniSapiens 2.0** demonstrates strong zero-shot transfer to OOD contexts!
 - Autism Detection, Speech Emotion Recognition, Social Media Sentiment, Sarcasm, Depression Recognition
 - Relevant for data-scarce behavioral contexts : Mental Health etc.

RQ3: How Can Behavioral Reasoning be Enhanced via Reasoning-RL?

vs. Baselines	Specificity			Coherence			Concision		
	Tie+Win	Tie	Win [†]	Tie+Win	Tie	Win [†]	Tie+Win	Tie	Win [†]
Qwen2.5-Omni-7B	80.6	35.3	66.0	90.0	24.0	85.3	99.3	13.3	99.3
Gemma 4-E4B (8B)	76.0	16.7	70.0	88.0	12.0	85.4	98.7	0.7	98.7
HumanOmniV2	76.7	14.0	70.4	87.3	9.3	85.7	100.0	0.0	100.0
OmniSapiens-7B-RL	82.5	40.6	66.1	87.9	20.9	83.2	98.6	8.0	98.6
Avg. Rates	78.9	26.6	68.5	88.3	16.5	85.1	99.2	5.5	99.2



Q: Based on the audio and video, which of the autism-related behaviors are present? {Definitions of behaviors are provided, i.e: Aggressive Behavior: ... Non-Responsiveness: ... }

You may respond with one or more behavior.

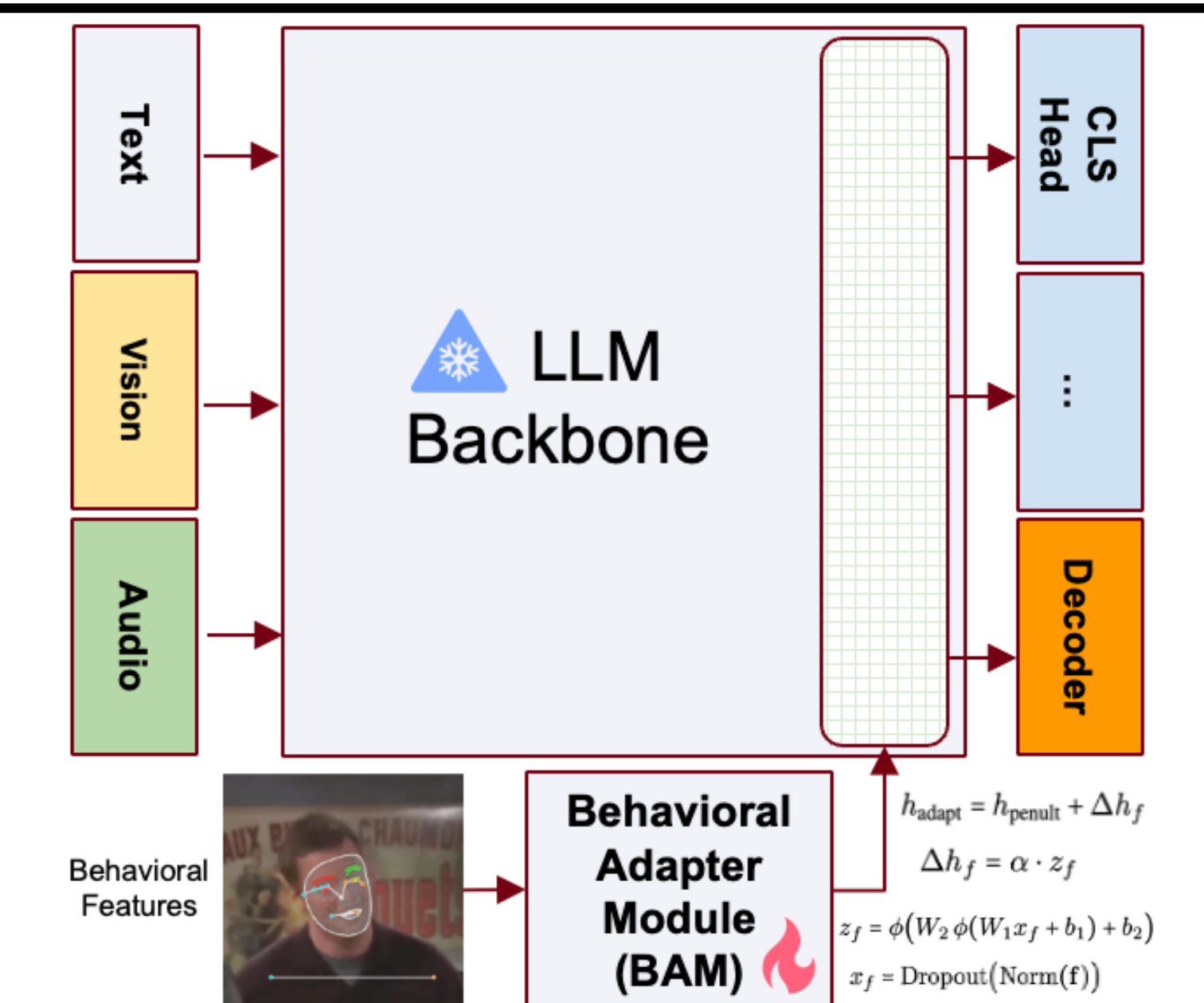
HumanOmniV2: Reasoning: He's standing and just swung his arms really fast and hard. That could definitely be seen as aggressive behavior... He's moving so fast too, looking quite agitated. This could be linked to hyperactivity, which might be part of sensory overresponsive behavior... He's sitting at the table, hands covering his face, looking distressed... maybe non-responsive to what's happening around him... Classification: Aggressive Behavior, Hyper- or Hyporeactivity to Sensory Input, Non-Responsiveness to Verbal Interaction ❌	OmniSapiens RL: Reasoning: The audio and video do not provide enough information to determine the presence of the autism-related behaviors listed. The background (label) is indicated as "no significant target behavior present" Classification: Background ❌	OmniSapiens 2.0: Reasoning: The children are hitting each other, which is aggressive behavior Classification: Aggressive Behavior ✅
--	--	--

- **OmniSapiens 2.0** resolves learning imbalances during training, coinciding with more interpretable reasoning traces!
 - More specific, coherent, concise reasoning traces
 - Greater reasoning consistency!

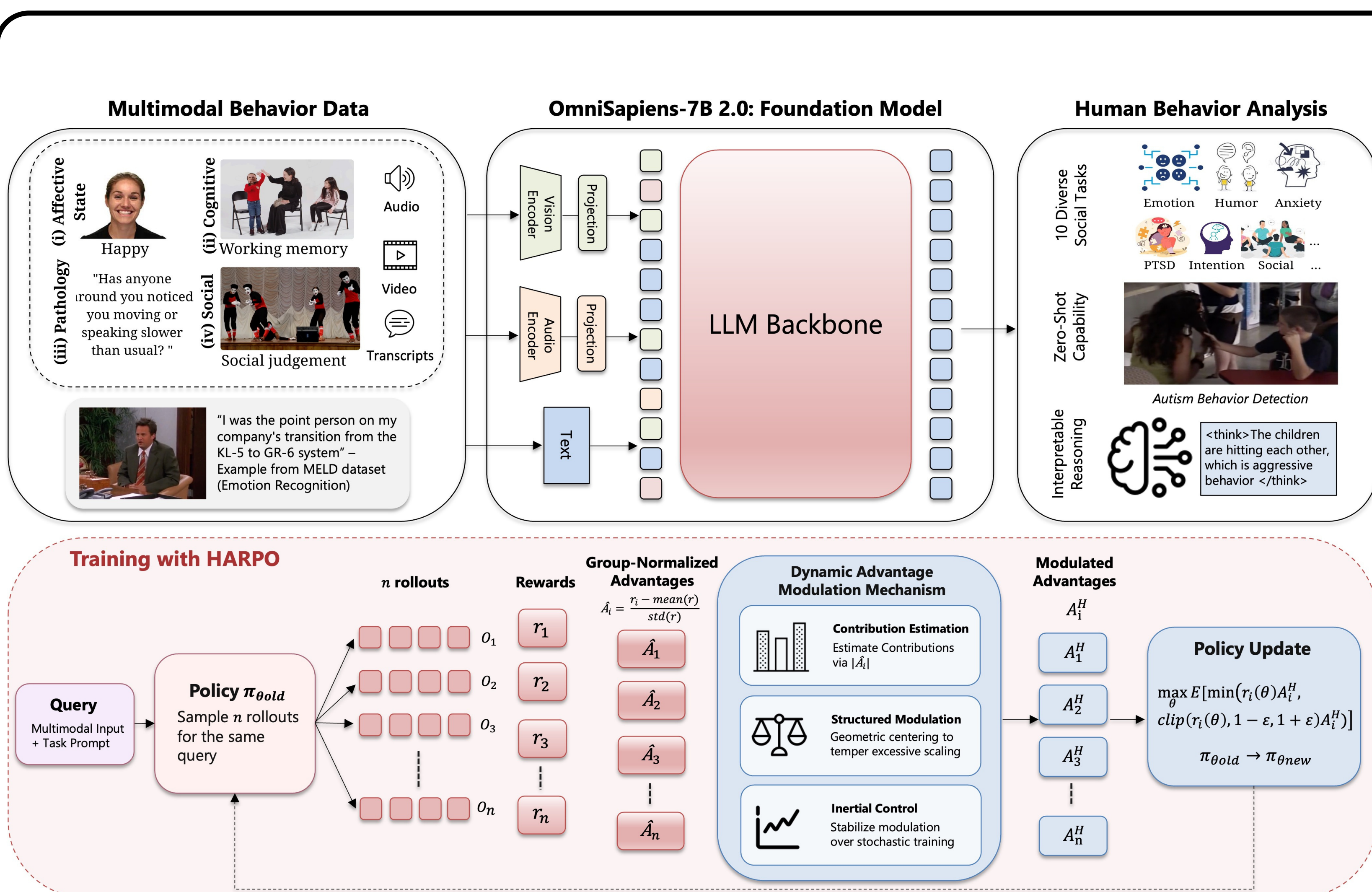
RQ4: Can Behavioral Foundation Models be Enhanced via Behavioral Adapter Modules?



BAM Analysis: The speaker briefly displays a split-second smile, signaling positive sentiment. While undetected by vanilla LLMs, Late-fusion via BAM improves the LLM's sensitivity to fine-grained signals.



OmniSapiens: New Developed Models, Architectures, Reasoning-RL



OmniSapiens-7B 2.0

- Heterogeneity-Aware RL (HARPO)
- Problem: Standard GRPO → dominated by dense-advantage tasks which creates learning imbalances
- Normalize task/sample contributions via geometric mean reference
- Apply multiplicative advantage reweighting
- Use EMA smoothing for stability