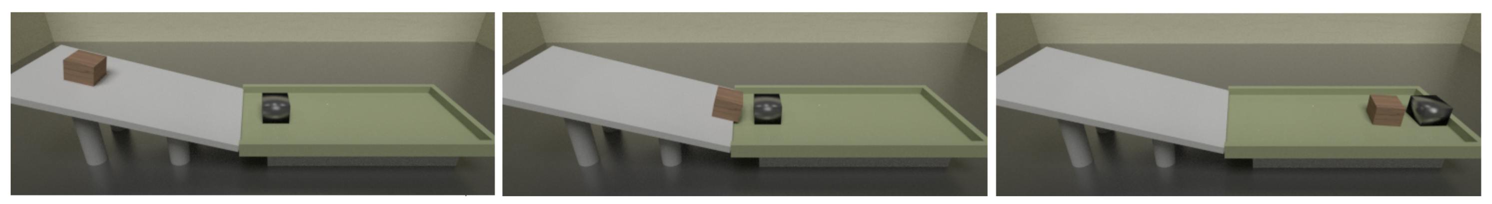
Real-time inference of physical properties in dynamic scenes

Kevin Smith[1,2], Mario Belledonne[1], Jiajun Wu[2,3], Josh Tenenbaum[1,2,3], Ilker Yildirim[1,2,4]

[1]Dept. of Brain and Cognitive Sciences, MIT | [2]Center for Brains, Minds, & Machines | [3]CSAIL, MIT | [4]Department of Psychology, Yale University

Motivation

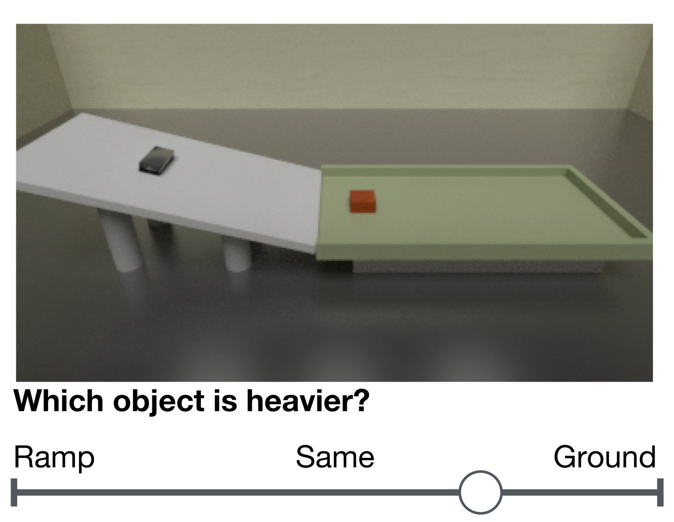
- Human scene understanding involves inferring latent causal properties – e.g., objects' mass
- These properties can be inferred in real-time from (1) texture, (2) object dynamics
- How do we do this "online" updating?
- Test via incongruencies in texture, dynamics



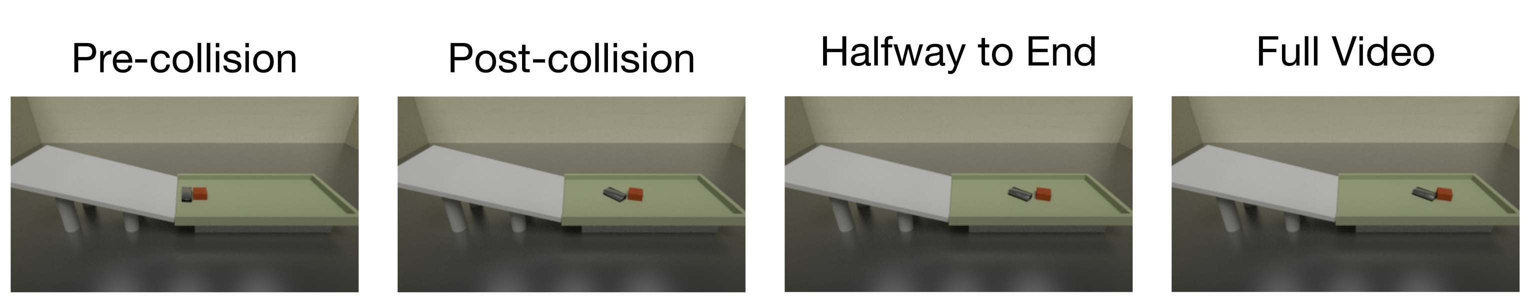
Experiment

Task

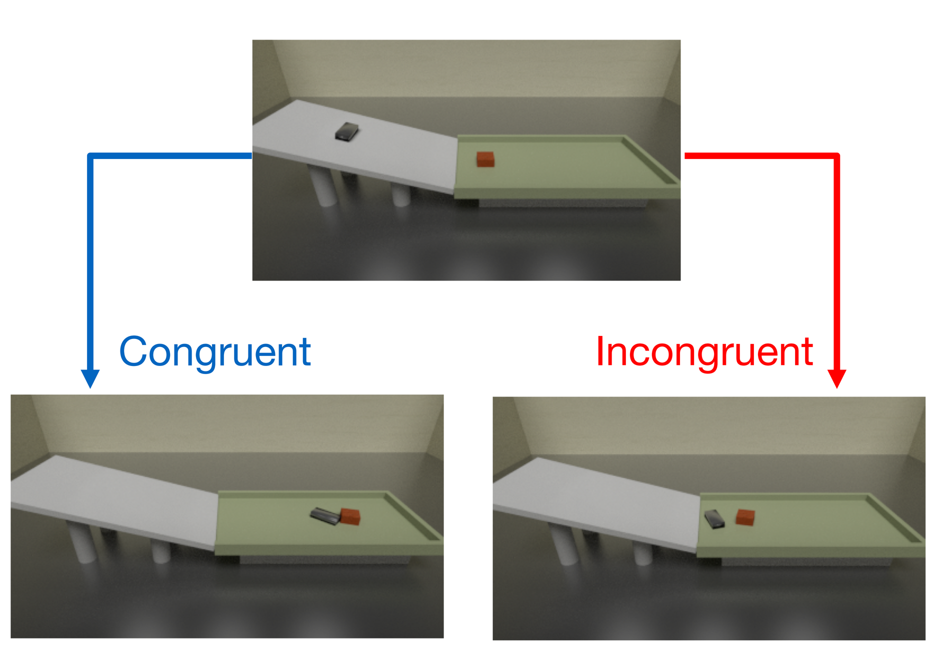
- Watch videos of dynamic collisions
- Judge relative masses



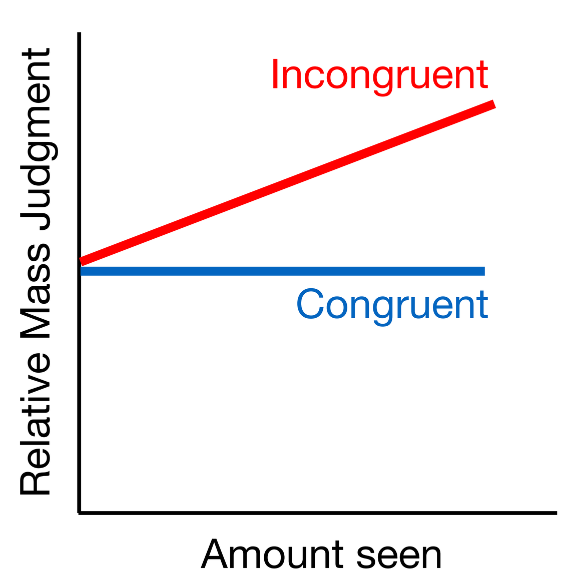
Four video cuts / scene



20% "incongruent" scenes



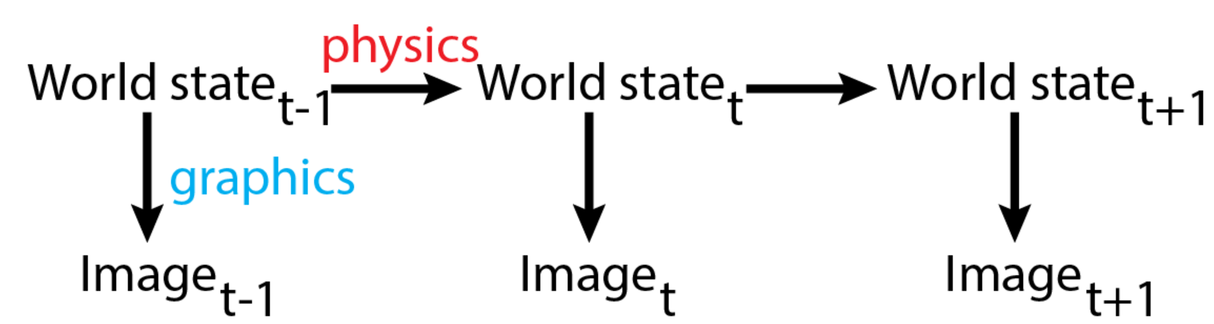
Hypothesis:



Dynamic Physical Inference

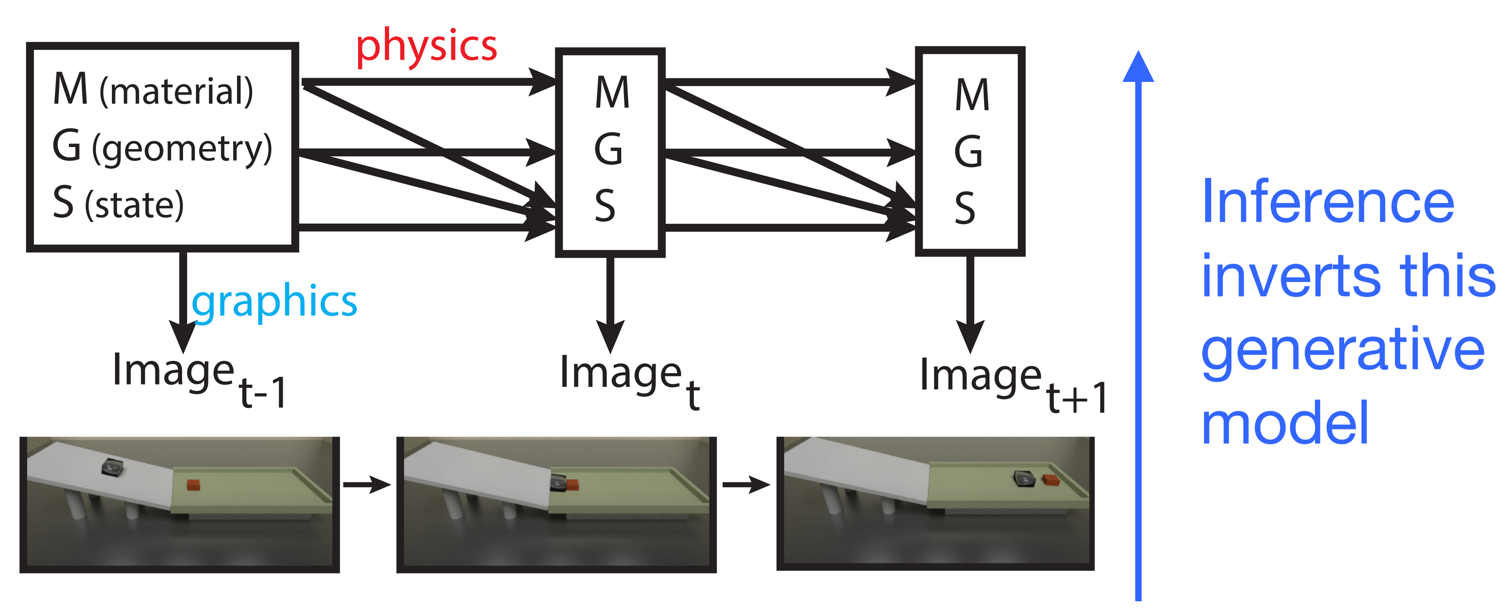
"Analysis by Synthesis:" reconstructing the scene that gives rise to sense inputs

Generative model



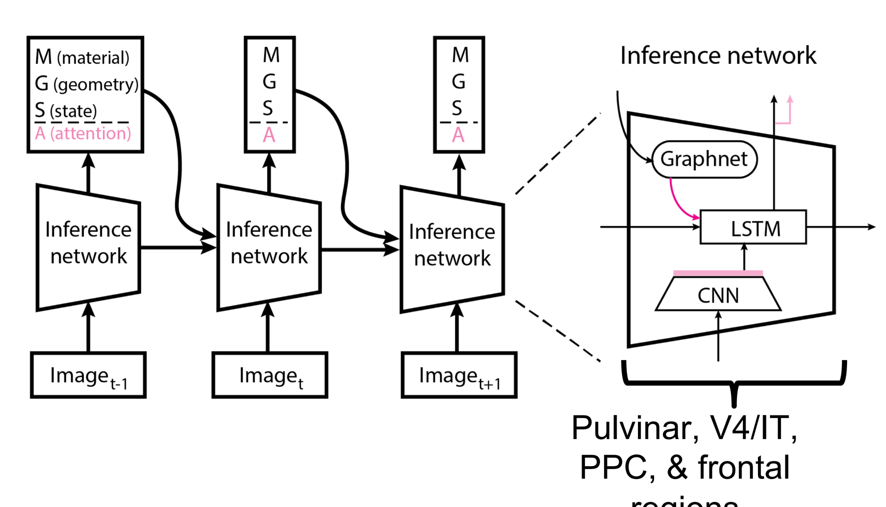
Battaglia, Hamrick, Tenenbaum (2013)

Inference problem

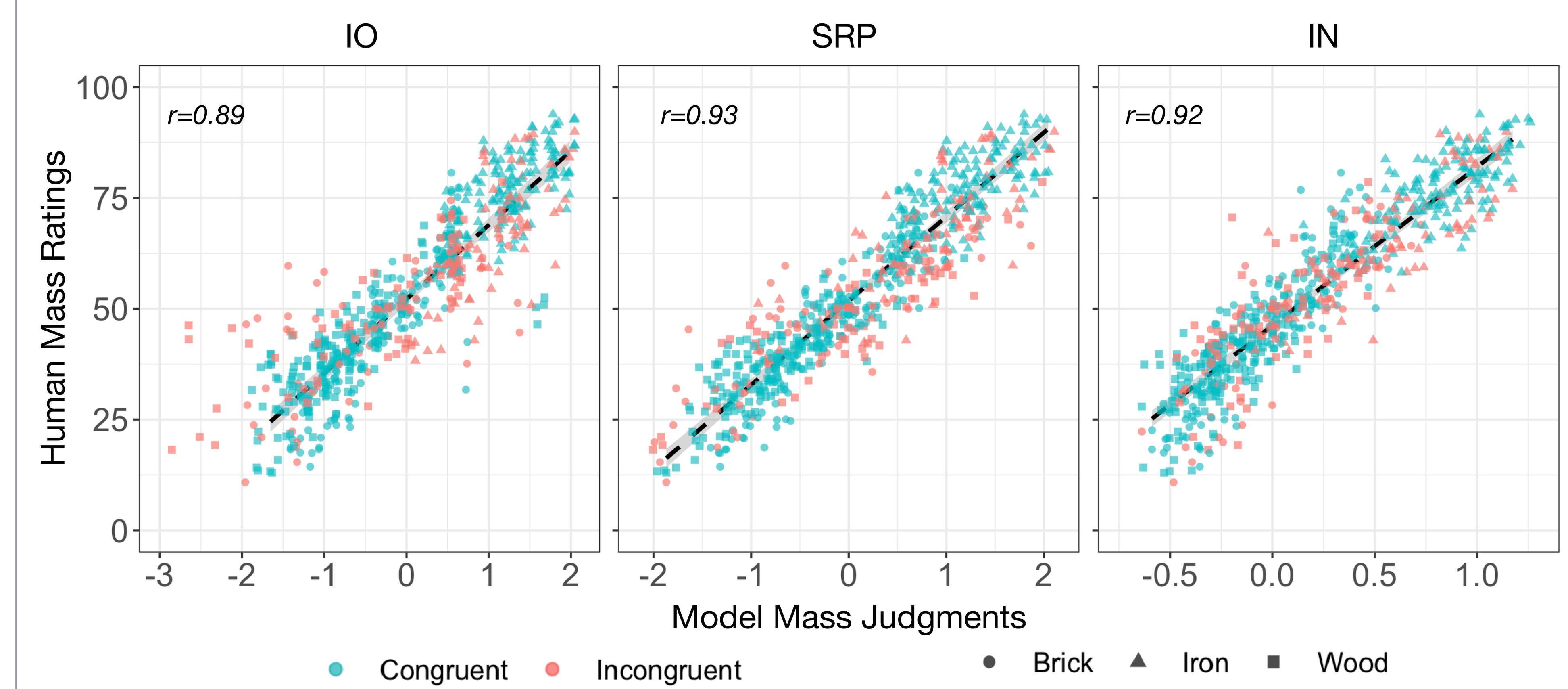


Inference Models

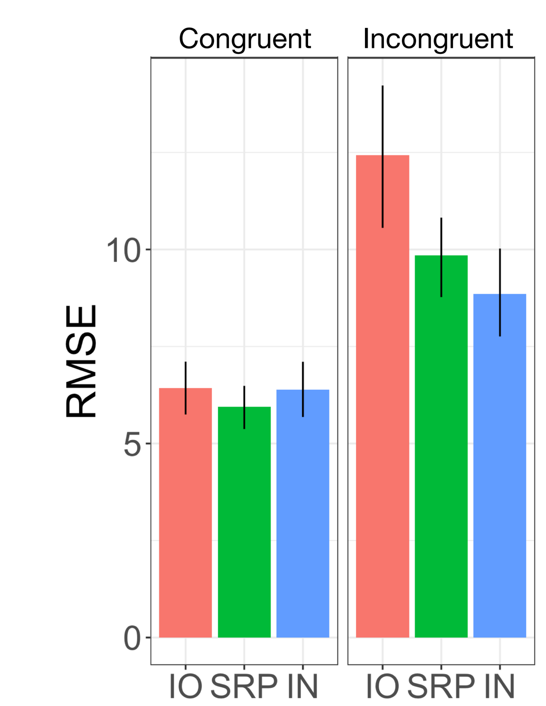
1. **Ideal Observer**
   - Exact inference at each time point
2. **Sequential Rational Process**
   - Approximation using Particle Filters
   - Resource constrained: 4 hypotheses at a time
3. **Inference Network**
   - Learns approximate inverse function
   - Plausible neural mapping?

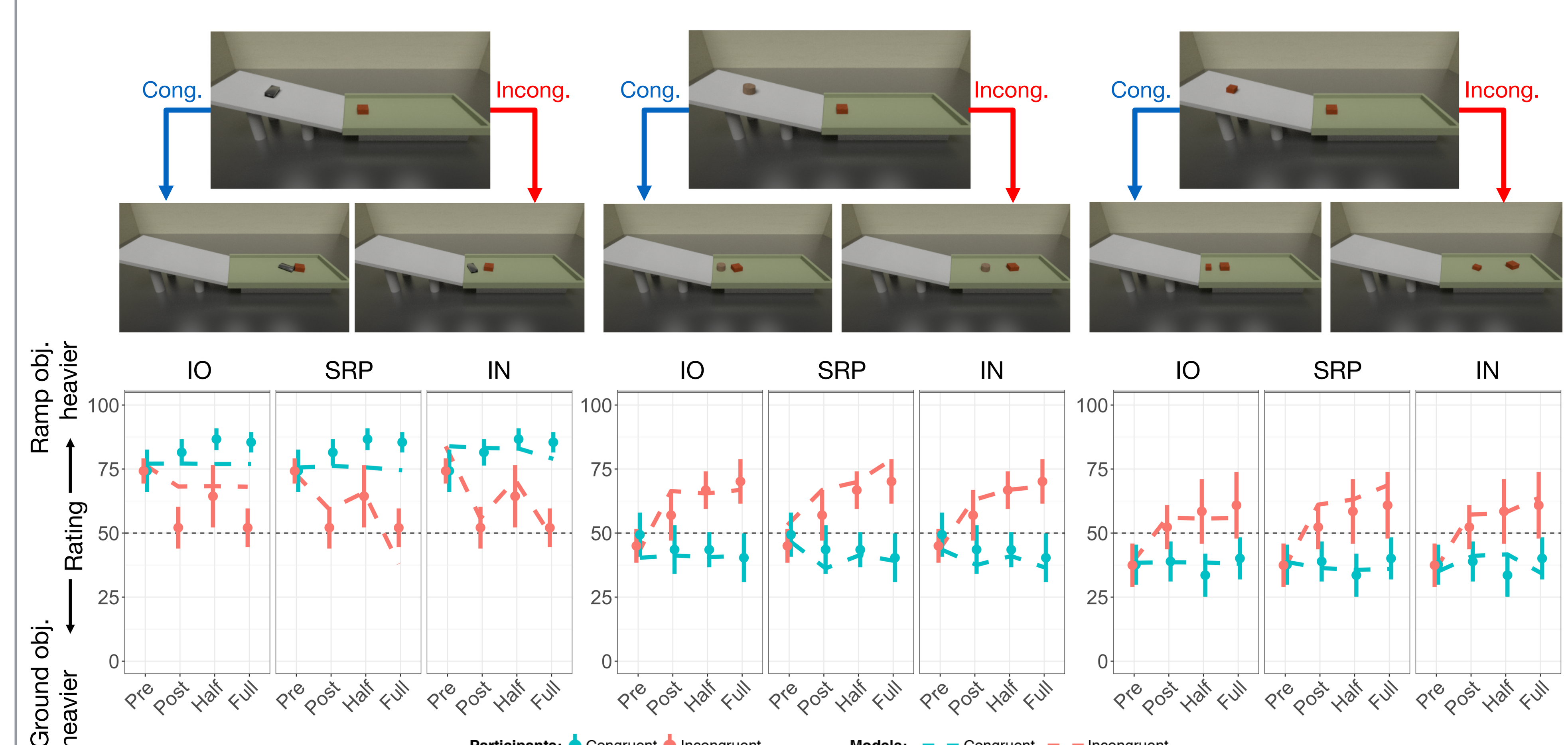


Results



- All models explain human judgments well overall
- But fit in IO is driven by congruent trials, whereas approximate models extend well to incongruent
- Approximate models capture dynamics of human judgment over time





Discussion

- Bayesian inference provides rich, causal explanations of sensory inputs supporting physical scene understanding
- Dynamic inference driven by efficient, approximate algorithms
- IN provides starting point for understanding physical inference in the brain

Acknowledgments

This material is based upon work supported by the Center for Minds, Brains, and Machines (CBMM), funded by NSF STC award CCF-1231216, and by a grant from Mitsubishi MELCO.