Perception as a Foundation for Common-Sense Theories of the World

Abdul-Rahim Deeb (adeeb1@jh.edu)
Department of Cognitive Science, Johns Hopkins University
Judith Fan (jefan@stanford.edu)
Department of Psychology, Stanford University
Shari Liu (sliu199@jhu.edu)
Department of Psychology, Johns Hopkins University
Kevin A Smith (k2smith@mit.edu)
Department of Brain and Cognitive Sciences, MIT

Keywords: Perception; Cognition; Intuitive Theories; Physical Reasoning; Social Reasoning

Overview

Humans have the remarkable ability to construct common-sense theories about the world through perception. From inferring the physical relationships between objects to understanding social interactions and predicting future events, perception plays a foundational role in shaping how we learn and reason. But how do systems for perception and higher-level cognitive theories interact? Are perception and common sense knowledge distinct, such that perception simply provides information for other cognitive systems to act on (Firestone & Scholl, 2016)? Or are they interlinked in ways that allow common-sense theories to affect what we perceive, or represent the world for other commonsense domains (Levin, Baker, & Banaji, 2016)?

By integrating insights from vision science, cognitive development, computational modeling, and neuroscience, we aim to uncover how perception supports and integrates with higher-order cognition. The symposium brings together researchers who investigate perception and cognition from diverse perspectives, offering insights into how perceptual systems shape not only our immediate understanding of the world but also the common-sense theories that guide our behaviors. These speakers will explore questions such as how human perception of social agents is grounded in physical reasoning (*Liu*), how visual abstractions are used to facilitate collaborative reasoning (*Fan*), how humans internalize physics through sensorimotor regularities in experience (*Deeb*), and how physical knowledge influences the way we perceive and represent the world (*Smith*). We will conclude with a moderated panel discussion aimed at synthesizing ideas, identifying open questions, and charting new directions for exploring the interplay between perception and cognition.

The Patchwork Approach: Toward a perceptual theory of intuitive physics

Abdul-Rahim Deeb

Human sensorimotor experience is rich with physical information, but it has been challenging to identify generalizable principles that explain how humans use experiential regularities to understand and predict the physical world. Here we present the Patchwork Approach: a method for modeling sensorimotor predictions based on experiential regularities, without the need for explicitly encoding physical laws. The Patchwork Approach leverages environmental regularities to enable humans to make perceptual predictions in a resource-rational way. While other recent methods offer powerful, data-driven approaches for learning physical laws (Piloto, Weinstein, Battaglia, & Botvinick, 2022), our approach emphasizes a more efficient use of prior sensory experiences to make predictions without the need for extensive computational resources.

Using data from previous studies (Deeb, Cesanek, & Domini, 2021; Deeb & Domini, 2024) on human perception of physical systems, the model outperforms traditional Newtonian models in capturing human prediction errors and interpolates across previously untested conditions. One test demonstrates its superiority in a projectile motion task, while another illustrates its ability to predict deflection angles in a collision event, even on untrained aiming angles. The Patchwork Approach provides valuable insights into how physical laws are internalized from the environment and used to guide perception and action and argues that perception provides foundational input to intuitive physics reasoning, a role that can complement and be extended to higher-level cognitive processes like those explained by the Intuitive Physics Engine (IPE).

Visual abstraction for planning and collaboration

Judith Fan

How do humans leverage abstract representations of visual inputs to solve complex problems together? In this talk, I will review our recent work examining how humans leverage visual abstraction to reason about physical assembly problems and coordinate to solve them together. Across several studies, we find converging evidence that people decompose physical structures in ways that make them both more tractable to think about and more efficient to talk about (Binder, Mattar, Kirsh, & Fan, 2023; Wong et al., 2022; McCarthy, Matejka, Willis, Fan, & Pu, 2024). These findings are compatible with the notion that the need to be able to coordinate on shared task representations under cognitive resource constraints can lead to strong alignment between the contents of visual scene representations and the organization of plans for achieving joint goals. More broadly, this work has implications for advancing our understanding of how visual computations relate to reasoning, communication, and action, as well as the development of AI systems that can effectively coordinate with humans through shared representations.

Perceiving and reasoning jointly about the social and physical world

Shari Liu

How do people make sense of other people, who are simultaneously psychological beings and physical objects? Across the cognitive sciences, researchers have studied theory of mind (making sense of other people's behaviors in terms of their mental states, or 'naive psychology') and physical reasoning (making sense of physical events in terms of their underlying mechanics and dynamics, or 'naive physics'), as two separate processes. In this talk, I will describe two key ways in which psychological reasoning depends on physical reasoning starting early in human development. First, we represent animate agents as objects who act on and in a physical world. Second, we use physical knowledge in order to make inferences about other minds, including what other people want, feel, and know, how hard they are trying, and how much danger they are in. We review research from developmental psychology (Gergely & Csibra, 2003; Liu, Ullman, Tenenbaum, & Spelke, 2017; Gjata, Ullman, Spelke, & Liu, 2022) and cognitive neuroscience (Pramod, Cohen, Tenenbaum, & Kanwisher, 2022; Liu, Lydic, Mei, & Saxe, 2024; Karakose-Akbiyik, Caramazza, & Wurm, 2023), which provides evidence for the intersection of these two systems, and Bayesian computational models of theory of mind (Baker, Jara-Ettinger, Saxe, & Tenenbaum, 2017; Jara-Ettinger, Gweon, Schulz, & Tenenbaum, 2016), which articulate a formal hypothesis about how these two systems work together (see also: Jara-Ettinger & Schachner, 2024; Shu, Peng, Zhu, & Lu, 2021). We propose that from infancy, we solve this 'commonsense mind-body problem' by dedicating two distinct but interconnected systems for perceiving and reasoning about ethereal minds and physical bodies, grounded in a shared representation of the physical world.

How physics influences the perception and representation of scenes

Kevin A Smith

How does our knowledge of physics affect what we perceive? Much of the research into intuitive physics uses the assumption that perception provides us with a noisy but relatively unbiased and complete representation of our environment, which acts as input to a probabilistic simulator (Smith et al., 2024). But recent findings have challenged that assumption, showing, for instance, that we often fail to notice when a block in a tower disappears from the scene (Ludwin-Peery, Bramley, Davis, & Gureckis, 2021) or do not account for the effect an object will have on the scene (Bass, Smith, Bonawitz, & Ullman, 2022).

Here I will argue that perception acts *in service* of physical reasoning, and so our intuitive physics impacts how we perceive and represent the world around us. First I will show how our mind selectively represents objects in the world that are relevant for our physical simulations (Chen, Allen, Cheyette, Tenenbaum, & Smith, 2023, *in prep*). I will then discuss ongoing work jointly modeling vision and physics that demonstrates how physical reasoning can help disambiguate perceptual uncertainty about objects' shapes and poses. Together this suggests a framework in which perception and intuitive theories work together to produce limited but powerful representations of the world around us.

References

Baker, C. L., Jara-Ettinger, J., Saxe, R., & Tenenbaum, J. B. (2017, March). Rational quantitative attribution of beliefs, desires and percepts in human mentalizing. *Nat. Hum. Behav.*, *1*(4), 1–10.

Bass, I., Smith, K. A., Bonawitz, E., & Ullman, T. D. (2022). Partial mental simulation explains fallacies in physical reasoning. *Cognitive Neuropsychology*, *38*(7-8), 413–424. doi: 10.1080/02643294.2022.2083950

Binder, F. J., Mattar, M. G., Kirsh, D., & Fan, J. E. (2023). Humans choose visual subgoals to reduce cognitive cost. In *Proceedings of the annual meeting of the cognitive science society* (Vol. 45).

Chen, T., Allen, K. R., Cheyette, S. J., Tenenbaum, J., & Smith, K. A. (2023). "Just In Time" representations for mental simulation in intuitive physics. In *Proceedings of the annual meeting of the cognitive science society* (Vol. 45).

Deeb, A.-R., Cesanek, E., & Domini, F. (2021). Newtonian predictions are integrated with sensory information in 3d motion perception. *Psychological Science*, *32*. doi: 10.1177/0956797620966785

Deeb, A.-R., & Domini, F. (2024). Embeddedness of earth's gravity in visual perception. *Journal of Vision*, *24*. doi: 10.1167/jov.24.11.4

Firestone, C., & Scholl, B. J. (2016). Cognition does not affect perception: Evaluating the evidence for "top-down" effects. *The Behavioral and brain sciences*, *39*, e229. doi: 10.1017/S0140525X15000965

Gergely, G., & Csibra, G. (2003, July). Teleological reasoning in infancy: the naıve theory of rational action. *Trends Cogn. Sci.*, *7*(7), 287–292.

Gjata, N. N., Ullman, T. D., Spelke, E. S., & Liu, S. (2022, July). What could go wrong: Adults and children calibrate predictions and explanations of others' actions based on relative reward and danger. *Cogn. Sci.*, *46*(7), e13163.

Jara-Ettinger, J., Gweon, H., Schulz, L. E., & Tenenbaum, J. B. (2016, August). The naïve utility calculus: Computational principles underlying commonsense psychology. *Trends Cogn. Sci.*, *20*(8), 589–604.

Jara-Ettinger, J., & Schachner, A. (2024, September). Traces of our past: The social representation of the physical world. *Curr. Dir. Psychol. Sci.*.

Karakose-Akbiyik, S., Caramazza, A., & Wurm, M. F. (2023, June). A shared neural code for the physics of actions and object events. *Nat. Commun.*, *14*(1), 3316.

Levin, D. T., Baker, L. J., & Banaji, M. R. (2016). Cognition can affect perception: Restating the evidence of a top-down effect. *The Behavioral and brain sciences*, *39*, e250. doi: 10.1017/S0140525X15002642

Liu, S., Lydic, K., Mei, L., & Saxe, R. (2024). Violations of physical and psychological expectations in the human adult brain. *Imaging Neuroscience*, *2*.

Liu, S., Ullman, T. D., Tenenbaum, J. B., & Spelke, E. S. (2017, November). Ten-month-old infants infer the value of goals from the costs of actions. *Science*, *358*(6366), 1038–1041.

Ludwin-Peery, E., Bramley, N., Davis, E., & Gureckis, T. (2021). Limits on Simulation Approaches in Intuitive Physics. *Cognitive Psychology*, *127*. doi: 10.31234/osf.io/xhzuc

McCarthy, W. P., Matejka, J., Willis, K. D., Fan, J. E., & Pu, Y. (2024). Communicating design intent using drawing and text. In *Proceedings of the 16th conference on creativity & cognition* (pp. 512–519).

Piloto, L., Weinstein, A., Battaglia, P., & Botvinick, P. (2022). Intuitive physics learning in a deep-learning model inspired by developmental psychology. *Nature Human Behaviour*, *6*. doi: 10.1038/s41562-022-01394-8

Pramod, R. T., Cohen, M. A., Tenenbaum, J. B., & Kanwisher, N. (2022, May). Invariant representation of physical stability in the human brain. *Elife*, *11*.

Shu, T., Peng, Y., Zhu, S.-C., & Lu, H. (2021, August). A unified psychological space for human perception of physical and social events. *Cogn. Psychol.*, *128*(101398), 101398.

Smith, K. A., Hamrick, J. B., Sanborn, A. N., Battaglia, P. W., Gerstenberg, T., Ullman, T. D., & Tenenbaum, J. B. (2024). Intuitive physics as probabilistic inference. In T. L. Griffiths, N. Chater, & J. B. Tenenbaum (Eds.), *Bayesian models of cognition : reverse engineering the mind.* Cambridge, MA: MIT Press.

Wong, C., McCarthy, W. P., Grand, G., Friedman, Y., Tenenbaum, J. B., Andreas, J., . . . Fan, J. E. (2022). Identifying concept libraries from language about object structure. *arXiv preprint arXiv:2205.05666*.