A simulation-heuristics dual-process model for intuitive physics

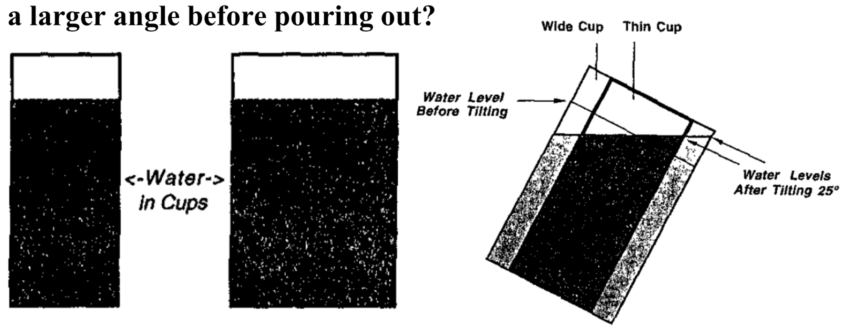
Shiqian Li* Peking University, BIGAI
Yuxi Ma* Peking University
Jiajun Yan* Peking University
Bo Dai BIGAI
Yujia Peng✉ Peking University, BIGAI
Chi Zhang✉ BIGAI
Yixin Zhu✉ Peking University

Introduction

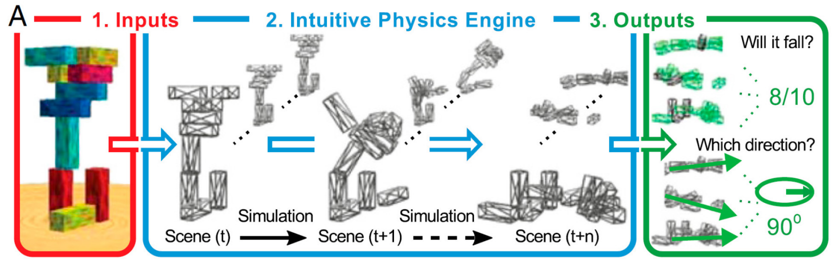
The computational mechanism of intuitive physics remains unclear.

People tend to be wrong when they answer quickly, but they are usually correct when they take the time to think carefully. (Schwartz & Black, 1999)

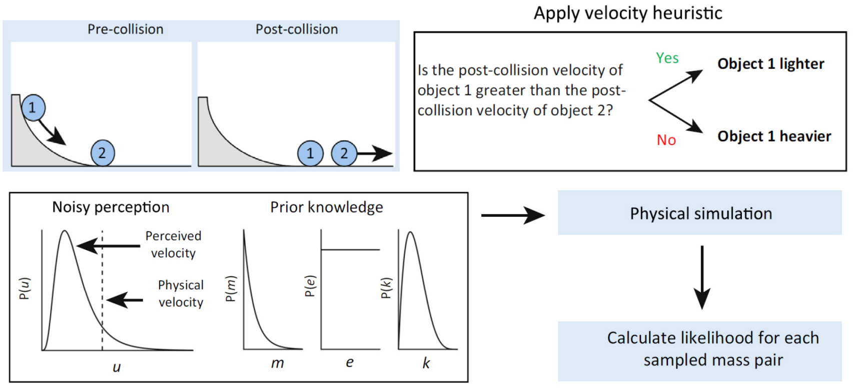
Which cup will need to be tilted with a larger angle before pouring out? **Answer: the thinner one**



Humans use “intuitive physics engine” to simulate scene stability, but use height as a heuristic to judge falling distance. (P. W. Battaglia et al., 2013)

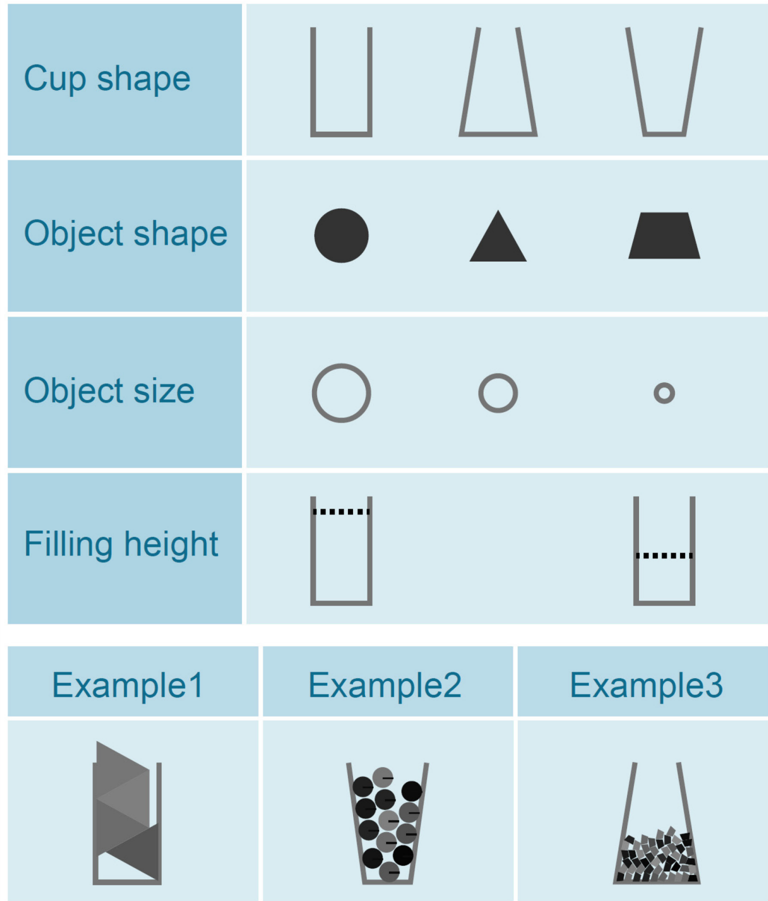


Humans apply velocity heuristic or use probabilistic simulation to judge the relative masses of colliding objects. (J. Kubricht et al., 2017)



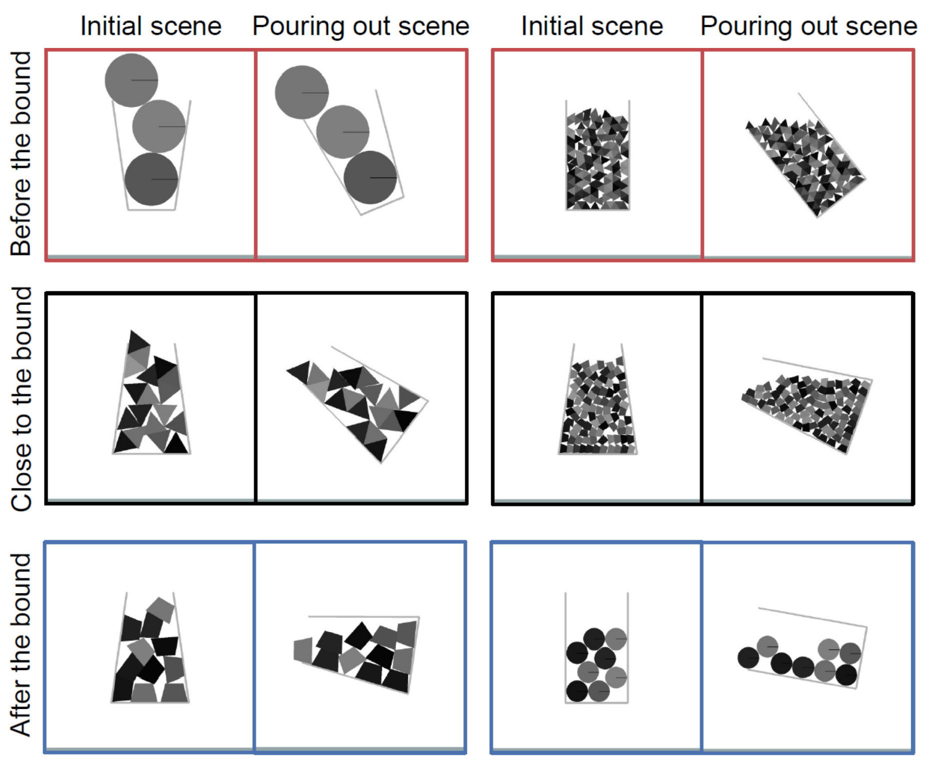
Question: Do humans consistently rely on mental simulation, or do they employ alternative heuristic strategies under certain conditions? How do they switch?

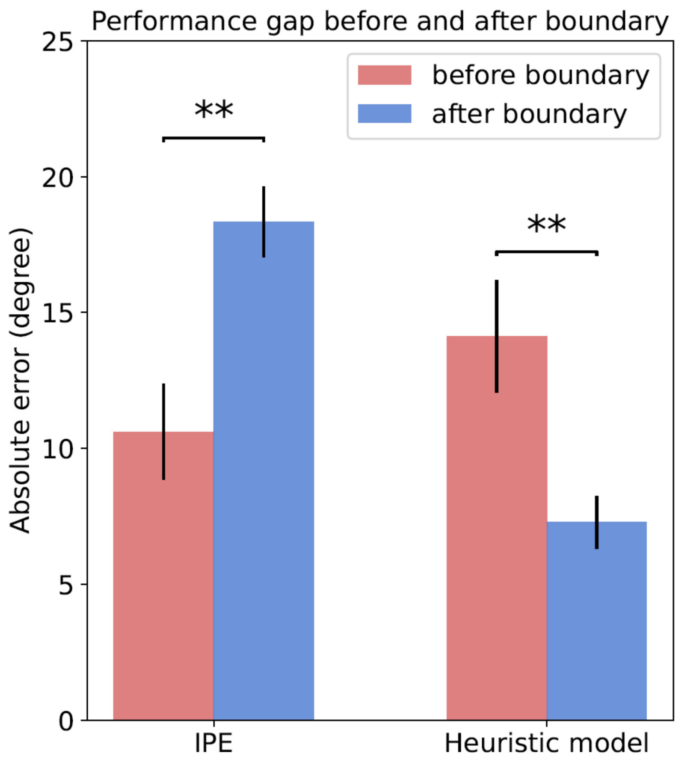
Experiments

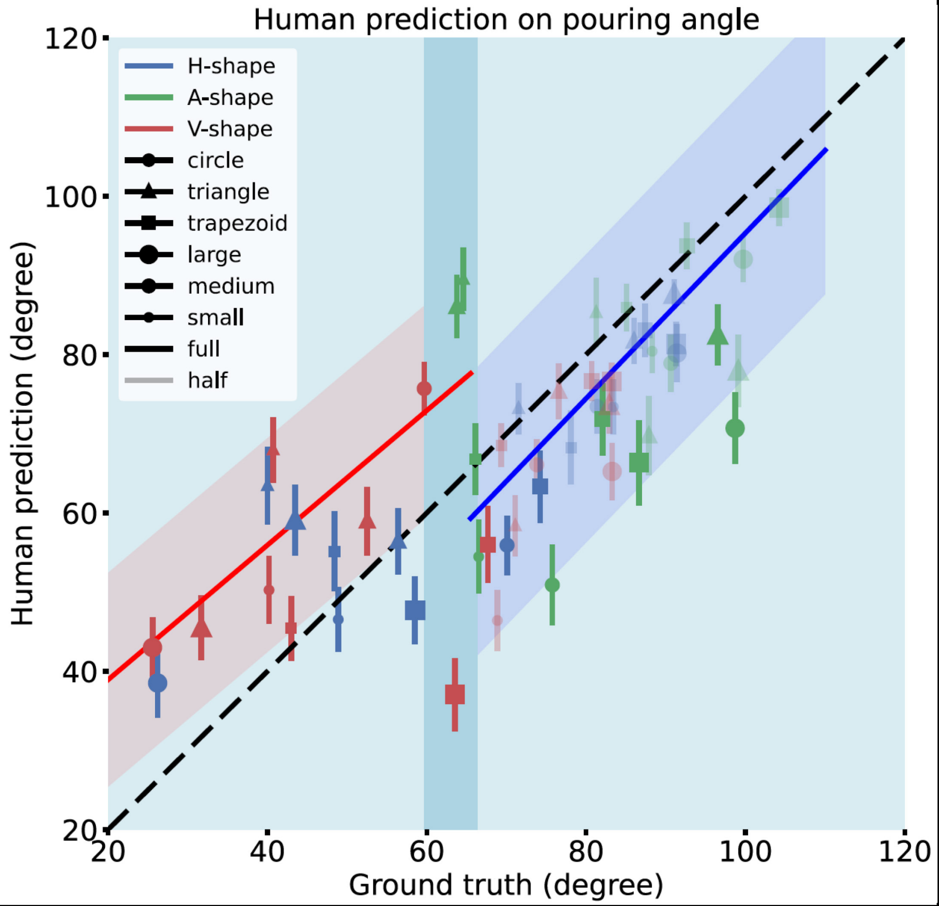


Task: Judge the tilting angle at which the cup begins to pour out marbles

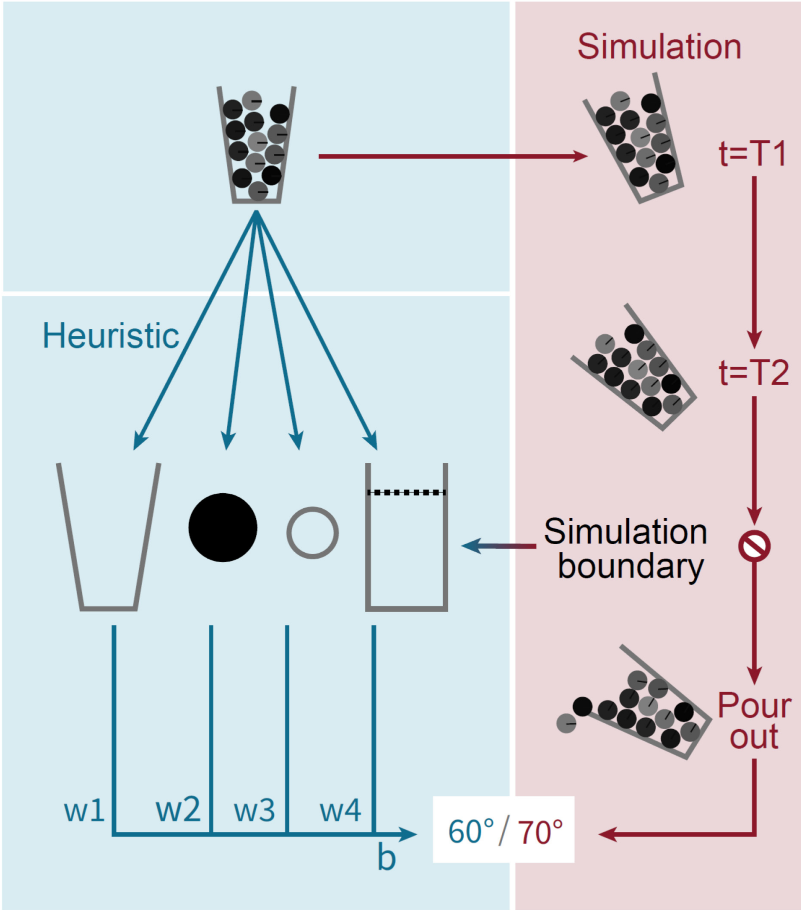
43 participants, 54 diverse scenarios, two error patterns (overestimation and underestimation)







Hypothesis



People have two mechanisms in physical reasoning, switching at a certain boundary:

Mechanism 1: probabilistic simulation

$$P(J|S_0, f_{0:T-1}) = \int_{S_{1:T}} P(J|S_{1:T})P(S_{1:T}|S_0, f_{0:T-1})dS_{1:T},$$

Mechanism 2: heuristic method

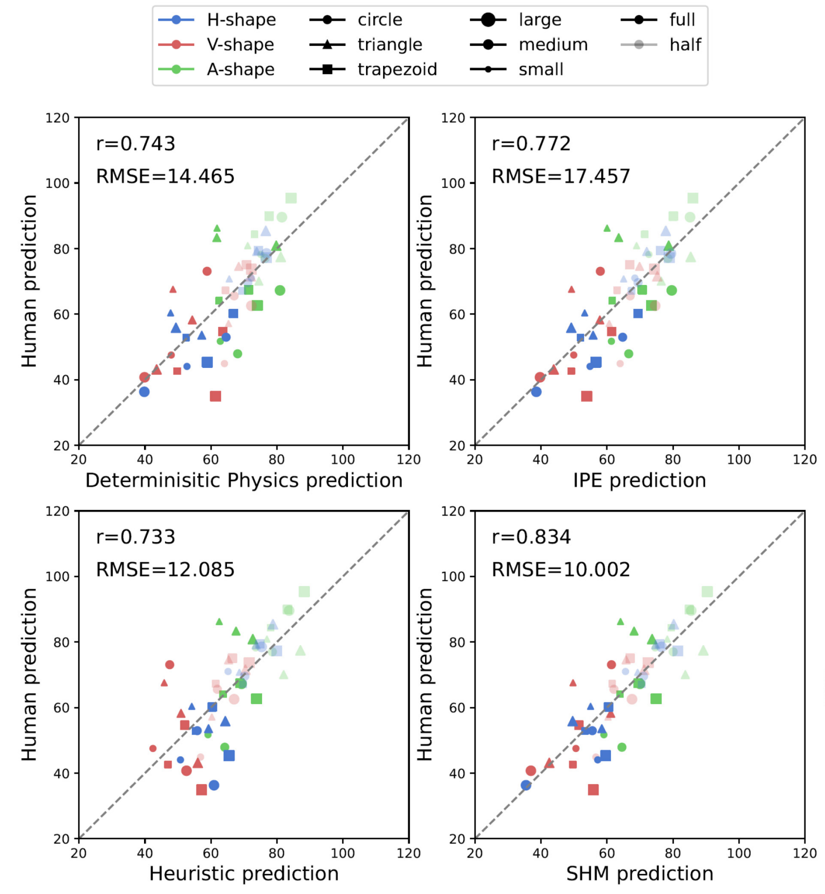
$$J = g(S_0^1, ..., S_0^n) = \sum_{i=1}^{n} \omega_i S_0^i + b,$$

Switch: when simulation reaches resource boundary.

$$\begin{cases} J = E_\epsilon[M(S_0;\epsilon)], & \text{if } T \leqslant \theta, \\ J = \sum_{i=1}^{n} \omega_i S_0^i + b, & \text{if } T > \theta \end{cases}.$$

where $S_{1:T}$ represents the sequence of all physical states from time steps 1 through T, and the integral averages over all possible trajectories of these states. Each state evolves as $S_{t+1} = \phi(S_t + \epsilon, f_t)$ with noise $\epsilon \sim \mathcal{N}(0, \sigma^2)$, and $\phi(\cdot)$ represents deterministic physical dynamics. We simplify the mapping from the initial state to the final judgment as $M(S_0; f, \epsilon)$.

Results



Our SHM aligns more precisely with human data, showing consistent predictive performance across diverse scenarios.