# Probabilistic Simulation Predicts Human Performance on Viscous Fluid-Pouring Problem

**James Kubricht**[⋆1] **kubricht@ucla.edu**
**Chenfanfu Jiang**[⋆3] **cffjiang@cs.ucla.edu**
**Yixin Zhu**[⋆2] **yixin.zhu@ucla.edu**
**Song-Chun Zhu**[2,3] **sczhu@stat.ucla.edu**
**Demetri Terzopoulos**[3] **dt@cs.ucla.edu**
**Hongjing Lu**[1,2] **hongjing@ucla.edu**

**Department of** [1] **Psychology,** [2] **Statistics, and** [3] **Computer Science**
⋆ Equal Contributors University of California, Los Angeles

## Abstract

The physical behavior of moving fluids is highly complex, yet people regularly interact with them with relative ease. To investigate how humans achieve this remarkable feat, we extended the classical water-pouring problem [1] to examine how humans consider physical properties of fluids (e.g., viscosity) and perceptual variables (e.g., volume) in a reasoning task. We found that humans do not rely on simple qualitative heuristics to reason about fluid dynamics. Computational results from a probabilistic simulation model can account for human sensitivity to hidden attributes and their performance on the water-pouring task. In contrast, non-simulation models based on statistical learning fail to fit human performance. The results in the present paper provide converging evidence supporting mental simulation in physical reasoning.

## 1 Introduction

Humans perceive and interact with many types of objects and substances on a daily basis; e.g., we regularly reason about fluid movement to predict when a filled container will spill. However, people perform significantly worse when asked to make explicit reasoning judgments in similar situations [1, 2]. Schwartz and Black [1] constructed a task requiring participants to choose which of two water-filled containers (one wider than the other) would need to be tilted farther before beginning to spill. Although participants failed the task in pencil-and-paper format, nearly all (95% of) the participants rotated a thinner, imaginary container (or a real, empty one) farther.

The *noisy Newton* framework for physical reasoning hypothesizes that inferences about dynamical systems can be generated by combining noisy perceptual inputs with the principles of classical (i.e., Newtonian) mechanics, given prior beliefs about represented variables [3–8]. In this framework, physical states are propagated forward in time using an *intuitive physics engine* given noisy perceptual inputs. The resulting predictions are queried and aggregated across simulations to determine the probability distribution of the associated human judgment. Bates et al. [3] extended the framework from physical scene understanding [4] to fluid dynamics using an *intuitive fluid engine* (IFE). Their particle-based IFE model using smoothed particle hydrodynamics (SPH; [9]) matched human judgments about final fluid states and provided a better quantitative fit than alternative models that did not employ simulation or account for physical uncertainty. The present study examined whether a particle-based IFE model can account for human judgments about the relative pour angle of two fluid-filled containers. The experiment reported here was inspired by previous empirical findings in water-pouring tasks (e.g., participants tilt containers filled with imagined molasses farther than those with an equal volume of water) indicating that people are able to take hidden attributes such as viscosity into account when making fluid-related judgments [1].

## 2 Experiment

### 2.1 Participants

A total of 152 participants were recruited from the Department of Psychology subject pool at the University of California, Los Angeles, and were compensated with course credit.

30th Conference on Neural Information Processing Systems (NIPS 2016), Barcelona, Spain.

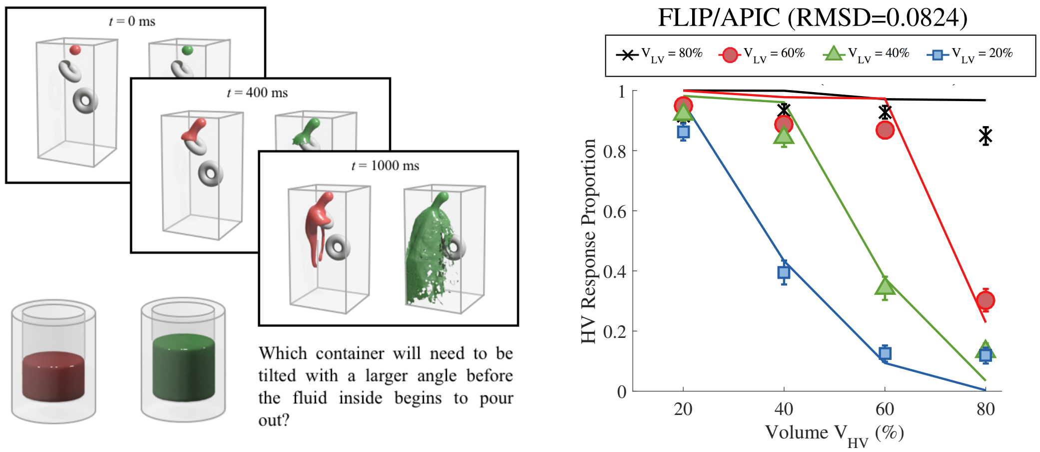


Figure 1: Flow demonstration illustration (left, top), judgment trial (left, bottom) and FLIP/APIC model predictions with perceptual noise (right). Separate lines indicate model predictions for each volume pair. Symbols indicate human response proportions.

### 2.2 Materials and procedure

Prior to the reasoning task, participants viewed animated demonstrations of a moderately viscous fluid's movement in two situations: (1) pouring over two torus-shaped obstructions and (2) spilling from a rotating container ($\omega = 22$ °/sec). The flow demonstration videos were presented to provide visual motion cues to inform participants' perceived viscosity [10]. Following the demonstration videos, two new fluids were introduced: one with low viscosity ($LV$; similar to water) and one with high viscosity ($HV$; similar to syrup). Participants viewed a side-by-side pouring video of both the $HV$ and $LV$ fluids before each judgment trial (see left panel of Fig. 1).

In the subsequent reasoning task, participants viewed a static image of two containers filled with the $LV$ and $HV$ fluids (see left panel of Fig. 1). Participants were instructed to assume that each container was tilted simultaneously at the same rate as observed earlier for the moderately viscous fluid in the demonstration. Participants were then asked to report "which container will need to be tilted with a larger angle before the fluid inside begins to pour out" and received no feedback following completion of each trial. The experiment manipulated the volume of the $LV$ and $HV$ fluids ($V_{LV}$ and $V_{HV}$, respectively) in each container across the values 20%, 40%, 60%, and 80%, representing the proportion of the container filled. Hence, the experiment consisted of 16 trials (reflecting all possible volume pairs) presented in a randomized order.

### 2.3 Human results

The symbols in Fig. 1 indicate human performance for all 16 fluid volume combinations. We examined whether humans rely on two candidate heuristics to make their judgments: (1) the container with less fluid will pour last (lesser-volume heuristic) and (2) the $HV$ fluid container will pour last (viscosity heuristic). For trials where $V_{HV}$ = 40%, 60%, and 80% and $V_{LV}$ = 20%, 40%, and 60% ($V_{LV} < V_{HV}$), the lesser-volume heuristic predicts $LV$ fluid responses. However, $HV$ response proportions for those trials were significantly greater thane zero ($t(151) = 9.92, 8.86, 8.10$, $p < .001$). Alternatively, the viscosity heuristic predicts unity $HV$ response proportions for the specified trials, which also disagrees with human response proportions. Thus, participants attended to latent fluid attributes (e.g., viscosity) *and* volume difference without employing a simple heuristic rule on either dimension when making their tilt angle judgments.

## 3 Models

### 3.1 Fluid simulation with physical dynamics

The velocity field of simulated fluids is determined using the Navier-Stokes equations. To numerically solve these partial differential equations, we adopt the Fluid Implicit Particle/Affine Particle in Cell (FLIP/APIC) method [11–13]. Unlike SPH, the FLIP/APIC method solves the Navier-Stokes

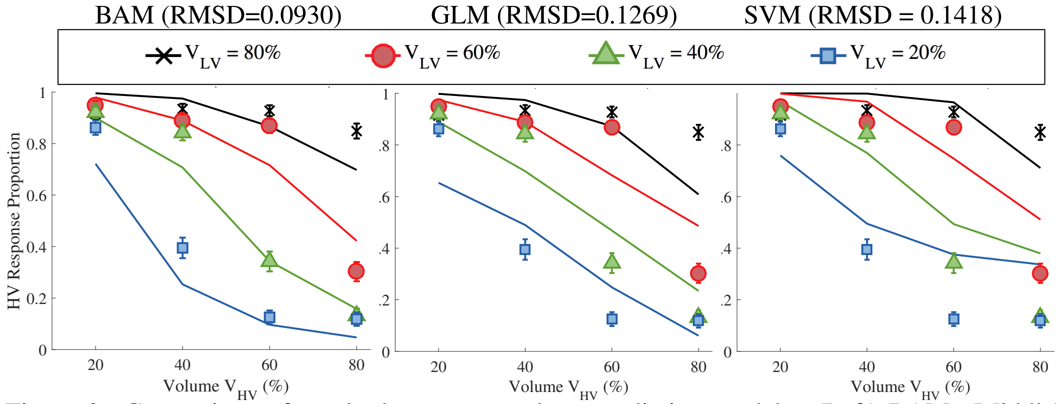


Figure 2: Comparison of results between our three prediction models: (Left) BAM, (Middle) Regression, and (Right) SVM with perceptual noise. Horizontal axes denote $HV$ fluid volume; vertical axes denote the predicted proportion of $HV$ fluid responses.

equations on a background Eulerian grid and maintains the benefits of particle-based methods due to its hybrid particle/grid nature. The presence of particles in the current model serves to facilitate visualization (important for flow demonstration animations) and the tracking of material properties. Since the FLIP/APIC method does not involve any stochastic processes, the output of each simulation is deterministic.

### 3.2 Intuitive fluid engine

The decisions directly derived from the deterministic FLIP/APIC fluid simulator given ground-truth parameters are binary judgments and thus cannot explain humans' probabilistic judgments. Inspired by the approach of Bates et al. [3] and the noisy Newton hypothesis [7], we combine the physical simulator of FLIP/APIC with noisy input variables to form the Intuitive Fluid Engine (IFE) model. Noisy volume is generated by adding an offset to its ground-truth value from a Gaussian distribution, whereas noisy viscosity is generated by adding a fixed amount of Gaussian noise on a logarithmic scale [7]. Each sample is then passed to the FLIP/APIC simulator to produce a binary decision; i.e., $LV$ or $HV$ fluid response. The IFE outputs the response distribution by aggregating the predictions and dividing the sum by the number of samples. Response proportions for the FLIP/APIC model are indicated in the right panel of Fig. 1.

We also compared human response proportions to a second simulation model which approximated the fluid as a set of rigid balls (see left panel of Fig. 2). In the ball approximation model (BAM), viscosity was approximated by damping the angular acceleration, $\alpha_B$, of each ball with magnitude proportional to its angular velocity, $\omega_B$, weighted by a stiffness parameter, $s$. Similar to noisy viscosity, noisy stiffness was generated by offsetting the median value with Gaussian noise on a logarithmic scale.

### 3.3 Non-simulation models

To examine whether simulation is necessary to account for how humans reason about fluid behavior, we compare the simulation model with two statistical learning methods—the generalized linear model (GLM) [14] and the support vector machine (SVM) [15] with perceptual noise. These models are purely data-driven and do not involve any explicit knowledge of physical laws. The selected features for these models include (i) the volumes of fluids in both containers, and (ii) the viscosity ratio between the $LV$ and $HV$ fluids.

### 3.4 Modeling results

We first compared how well different computational models account for human performance for the 16 trials. The right panel of Fig. 1 depicts results from the FLIP/APIC model and Fig. 2 depicts results from the BAM, GLM, and SVM models with perceptual noise. Human judgments and model predictions were highly correlated ($r(14) = 0.9954, 0.9663, 0.9488,$ and $0.9251$, respectively). Compared to the purely data-driven models (i.e., the GLM and SVM models), the simulation-based IFE and ball approximation models provide better approximations to human judgments in the viscous fluid-pouring task. These results support the role of simulation as a potential mental model that supports human inference in physical reasoning tasks.

## 4 Discussion

Our results from the viscous fluid-pouring task agree with the findings of Bates et al. [3] in that (1) our probabilistic, simulation-based IFE and ball approximation models outperformed two non-simulation models (SVM and GLM) and (2) people naturally attend to latent attributes (e.g., viscosity) when reasoning about fluid states. Participants' performance in our reasoning task suggests representation of physical quantities that extends beyond qualitative or symbolic understanding. While human results are generally consistent with physics-based simulation models coupled with noisy input variables, there remain discrepancies between model predictions and human judgments. Hence, future research should aim to address whether humans *simulate* fluid movements using mental models that accord to physical laws or *emulate* fluid dynamics by drawing on their everyday interactions with liquids across diverse physical situations [16]. Moreover, the accuracy of our ball approximation model suggests that people may simulate approximated representations of substances using principles of rigid body mechanics rather than fluid dynamics.

## References


[1] Schwartz, D. L. and Black, T. (1999). Inferences through imagined actions: knowing by simulated doing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 25(1):116.

[2] McAfee, E. A. and Proffitt, D. R. (1991). Understanding the surface orientation of liquids. *Cognitive Psychology*, 23(3):483–514.

[3] Bates, C. J., Yildirim, I., Tenenbaum, J. B., and Battaglia, P. W. (2015). Humans predict liquid dynamics using probabilistic simulation. In *Proceedings of the 37th annual conference of the cognitive science society*.

[4] Battaglia, P. W., Hamrick, J. B., and Tenenbaum, J. B. (2013). Simulation as an engine of physical scene understanding. *Proceedings of the National Academy of Sciences*, 110(45):18327–18332.

[5] Gerstenberg, T., Goodman, N. D., Lagnado, D. A., and Tenenbaum, J. B. (2015). How, whether, why: Causal judgments as counterfactual contrasts. In *Proceedings of the 37th annual conference of the cognitive science society*.

[6] Sanborn, A. N. (2014). Testing bayesian and heuristic predictions of mass judgments of colliding objects. *Frontiers in psychology*, 5.

[7] Sanborn, A. N., Mansinghka, V. K., and Griffiths, T. L. (2013). Reconciling intuitive physics and newtonian mechanics for colliding objects. *Psychological review*, 120(2):411.

[8] Smith, K. A. and Vul, E. (2013). Sources of uncertainty in intuitive physics. *Topics in cognitive science*, 5(1):185–199.

[9] Monaghan, J. J. (1992). Smoothed particle hydrodynamics. *Annual review of astronomy and astrophysics*, 30:543–574.

[10] Kawabe, T., Maruya, K., Fleming, R. W., and Nishida, S. (2015). Seeing liquids from visual motion. *Vision research*, 109:125–138.

[11] Zhu, Y. and Bridson, R. (2005). Animating sand as a fluid. *ACM Transactions on Graphics (TOG)*, 24(3):965–972.

[12] Jiang, C., Schroeder, C., Selle, A., Teran, J., and Stomakhin, A. (2015). The affine particle-in-cell method. *ACM Transactions on Graphics (TOG)*, 34(4):51.

[13] Jiang, C. (2015). The material point method for the physics-based simulation of solids and fluids. *PhD dissertation, University of California, Los Angeles*.

[14] McCullagh, P. (1984). Generalized linear models. *European Journal of Operational Research*, 16(3):285–292.

[15] Cortes, C. and Vapnik, V. (1995). Support-vector networks. *Machine learning*, 20(3):273–297.

[16] Grzeszczuk, R., Terzopoulos, D., and Hinton, G. (1998). Neuroanimator: Fast neural network emulation and control of physics-based models. In *Proceedings of the 25th annual conference on computer graphics and interactive techniques*.