Probabilistic Simulation Predicts Human Performance on Viscous Fluid-Pouring Problem

James Kubricht[⋆1]
kubricht@ucla.edu

Chenfanfu Jiang[⋆2]
cffjiang@cs.ucla.edu

Yixin Zhu[⋆3]
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

[1] Department of Psychology [2] Department of Statistics [3] Department of Computer Science

[⋆] Equal Contributors University of California, Los Angeles

Abstract

The physical behavior of moving fluids is highly complex, yet people are able to interact with them in their everyday lives with relative ease. To investigate how humans achieve this remarkable ability, the present study extended the classical water-pouring problem (Schwartz & Black, 1999) to examine how humans take into consideration 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. Instead, they rely on the perceived viscosity and fluid volume to make quantitative judgments. Computational results from a probabilistic simulation model can account for human sensitivity to hidden attributes, such as viscosity, 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, in addition to developing a set of experimental conditions that rectify the dissociation between explicit prediction and tacit judgment through the use of mental simulation strategies.

Keywords: Intuitive physics; mental simulation; animation; reasoning

Introduction

Imagine that you are preparing to pour pancake batter onto a griddle. To pour the correct amount, you must decide where to hold the container, at what angle, and for how long. We encounter similar situations frequently in our daily lives when interacting with viscous fluids ranging from water to honey, and with different volumes, contained in receptacles of various shapes and sizes.

In the majority of these situations, we are able to reason about fluid-related physical processes so as to implicitly predict how far a filled container can be tilted before the fluid inside begins to spill over its rim. However, people perform significantly worse when asked to make explicit reasoning judgments in similar situations (McAfee & Proffitt, 1991; Schwartz & Black, 1999). In the well-known Piagetian water level task (WLT; Howard, 1978), participants receive instructions to draw the water level at indicated locations on the inside of tilted containers. Surprisingly, about 40% of adults predict water levels that deviate from the horizontal by 5 degrees or more (e.g., McAfee & Proffitt, 1991). Schwartz and Black (1999) modified the WLT to include two containers, one wider than the other. The investigators asked participants to judge which container would need to be tilted farther before the water inside begins to pour out. Only 34% of the participants correctly reported that the thinner container would need to be tilted farther than the wider one. However, when instructed to complete the task by closing their eyes and imagining the same situation, nearly all (95% of) the participants rotated a thinner, imaginary container (or a real, empty one) farther. These findings suggest that people are able to reason successfully about relative pour angles by mentally simulating the tilting event. An apparent contrast in human performance between an explicit reasoning task and a simulated-doing task has also been found in people's inferences about the trajectories of falling objects (Krist, Fieberg, & Wilkening, 1993; K. Smith, Battaglia, & Vul, 2013). Thus, empirical findings in the literature of physical reasoning suggest that people employ both explicit knowledge about physical rules *and* and mental simulation when making inferences (Hegarty, 2004).

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 (Bates, Yildirim, Tenenbaum, & Battaglia, 2015; Battaglia, Hamrick, & Tenenbaum, 2013; Gerstenberg, Goodman, Lagnado, & Tenenbaum, 2015; Sanborn, 2014; Sanborn, Mansinghka, & Griffiths, 2013; K. A. Smith & Vul, 2013). In this framework, the locations, motions and physical attributes of objects are sampled from noisy distributions and propagated forward in time using an *intuitive physics engine*. The resulting predictions are queried and averaged across simulations to determine the probability of the associated human judgment. Bates et al. (2015) extended the framework from physical scene understanding (e.g., Battaglia et al., 2013) to fluid dynamics using an *intuitive fluid engine* (IFE), where future fluid states are approximated by probabilistic simulation via a Smoothed Particle Hydrodynamics (SPH) method (Monaghan, 1992). The particle-based IFE model 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 aims to determine whether a particle-based IFE model coupled with noisy input variables can account for human judgments about the relative pour angle of two containers filled with fluids differing in their volume and viscosity. 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, suggesting that people are able to take physical attributes such as viscosity into account when making fluid-related judgments (Schwartz & Black, 1999). Bates et al. (2015) also found that their participants' judgments were sensitive to latent attributes of the fluid (e.g., stickiness and viscosity).

To quantify the extent that humans employ their perceived viscosity of fluids in subsequent reasoning tasks, we utilized a recent development in graphical fluid simulation (Bridson, 2008; Jiang, Schroeder, Selle, Teran, & Stomakhin, 2015) to simulate the dynamic behavior of fluids in vivid animations. Previous work has shown that realistic animations can facilitate representation of *dynamic* physical situations (Tversky, Morrison, & Betrancourt, 2002). Furthermore, recent research on human visual recognition indicates that latent attributes of fluids (e.g., viscosity) are primarily perceived from visual motion cues (Kawabe, Maruya, Fleming, & Nishida, 2015), so displaying realistic fluid movement is needed to provide the input of key physical properties that enable mental fluid simulations. The present study, which uses a modification of Schwartz and Black's (1999) water-pouring problem coupled with advanced techniques in computer graphics, aims to test the hypothesis that animated demonstrations of flow behavior facilitate inference of latent fluid attributes, which inform mental simulations and enhance performance in subsequent reasoning tasks.

Experiment

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.

Materials and Procedure Prior to the reasoning task, participants viewed animated demonstrations of the movement of a moderately viscous fluid in two situations. The fluid used in the demonstrations was colored orange and was not observed in the judgment task. In the first demonstration, the fluid pours over two torus-shaped obstructions in a video looped three times and lasting for 11.5 seconds. The flow demonstration videos were presented to provide visual motion cues to inform participants' perceived viscosity. Following the flow demonstration, participants viewed a video of a cylindrical container filled with the same orange fluid tilting at a constant angular rate ($\omega = 22$ deg$\cdot s^{-1}$; see Fig. 1) from the upright orientation of the container and moving towards the horizontal. The video was looped three times for a duration of 14.7 seconds.

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 molasses). The *LV* and *HV* fluids were colored either red or green (counterbalanced across subjects). As shown in the top panel of Fig. 1, participants viewed a flow video of both the *HV* and *LV* fluids (looped three times) for a duration of 11.5 seconds before each judgment trial. The two flow videos were presented side by side for comparison, and the relative position of each fluid was counterbalanced across subjects. The *LV* and *HV* fluids were selected to readily distinguish the two fluids based on their perceived viscosities, which were inferred from visual motion cues in the flow videos (Kawabe et al., 2015).

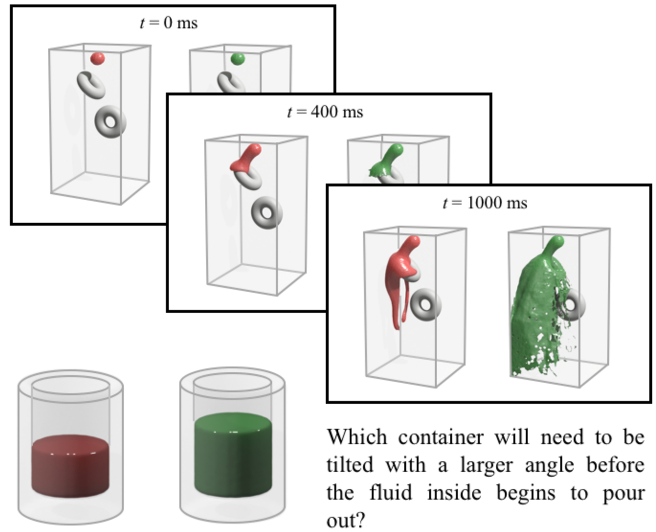

Figure 1: Illustration of flow demonstration video and judgment trial. (Top) Sample frames from the *HV* (red) and *LV* (green) flow video. (Bottom) Tilt judgment trial, where $V_{HV} = 40\%$ and $V_{LV} = 60\%$.

In the subsequent reasoning task, participants viewed a static image of two containers side by side filled with the *LV* and *HV* fluids (see bottom panel of Fig. 1). Participants were instructed to assume that each container was tilted simultaneously in the same way as observed earlier for the orange fluid in the tilting demonstration. They were informed that both containers were tilted at the same rate, and were provided with the quantity of fluid in each container. 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 presented in a randomized order, including all possible volume pairs between the *LV* and *HV* fluids. The experiment lasted approximately 10 minutes.

Human Results

Fig. 2 depicts human performance for all 16 fluid volume combinations. To assess the relationship between *HV* fluid volume and human judgments, we performed a logistic regression analysis and compared estimated coefficients for each participant across *LV* fluid volume conditions. We found that coefficients for each participant varied significantly between V_{LV} conditions ($F(3,149) = 113.89, p < .001$). In the highest *LV* fluid volume condition ($V_{LV} = 80\%$), V_{HV} had a

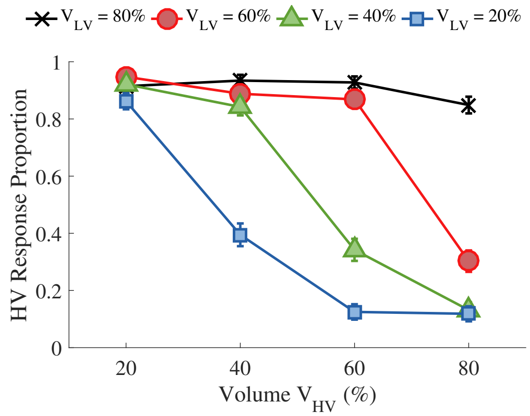

Figure 2: Human *HV* response proportions for all experimental conditions. The volume of the *HV* fluid (V_{HV}) is plotted on the horizontal axis, and separate lines indicate the four possible volumes for the *LV* fluid (V_{LV}).

minor impact on participants' responses ($\bar{\beta}_{80} = .04, \sigma_{\beta_{80}} = .25$) relative to the lowest *LV* fluid volume condition ($V_{HV} = 20\%$; $\bar{\beta}_{20} = .57, \sigma_{\beta_{20}} = .43$). These results demonstrate that participants' responses were increasingly sensitive to *HV* fluid volume for the greater V_{LV} conditions.

Next, we examined whether humans rely on heuristic-based reasoning to make their judgments. One potential heuristic is that given two containers filled with different volumes of each fluid, the container with lesser fluid volume requires a greater rotation before beginning to pour. While participants consistently adhered to this rule for trials where $V_{HV} < V_{LV}$, their judgments for each of the $V_{LV} < V_{HV}$ trials did not accord to the same heuristic. For example, in trials where V_{HV} = 40%, 60%, and 80% and V_{LV} = 20%, 40%, and 60%, respectively, the lesser-volume heuristic predicts *LV* fluid responses. However, *HV* response proportions for those trials were significantly greater thane zero ($t(151) = 9.92, 8.86, 8.10$, $p < .001$). A second potential heuristic is to always choose the *HV* fluid as requiring a greater rotation since it moves slower than the *LV* fluid. The above three cases also disagreed with this heuristic since the rule would predict unity response proportions. In summary, response proportions in the specified trials reveals that participants attended to latent fluid attributes (e.g., viscosity) *and* volume difference when making their tilt angle judgments (see Fig. 3).

Models

Fluid Simulation with Physical Dynamics

The simulation of incompressible flows through numerical evaluation of physical equations has become one of the most significant topics in computer graphics and mechanical engineering. The velocity field of simulated fluids is determined according to the constraints specified in the Navier-Stokes equations:

$$\frac{\partial \boldsymbol{v}}{\partial t} + \boldsymbol{v} \cdot \nabla \boldsymbol{v} + \frac{1}{\rho} \nabla p = \boldsymbol{g} + \mu \nabla \cdot \nabla \boldsymbol{v}, \tag{1}$$

$$\nabla \cdot \boldsymbol{v} = 0, \tag{2}$$

where $\boldsymbol{v}$ is the velocity, ρ is the density, p is the pressure, $\boldsymbol{g}$ is the gravitational acceleration (approximately 9.8 m s^{-2}), and μ is the viscosity. Equation (1) is called the momentum equation—it is simply Newton's second law (i.e., $F = ma$) applied to fluid dynamics. Equation (2) is the incompressibility constraint on fluid velocity, where the null divergence of the velocity field corresponds to constant density within volumetric regions. Most liquids need to satisfy this constraint in order to maintain constant volume while moving.

To numerically solve these partial differential equations, we adopt the Fluid Implicit Particle/Affine Particle in Cell (FLIP/APIC) method (Zhu & Bridson, 2005; Jiang et al., 2015; Jiang, 2015), which has become standard in physics-based simulation calculations due to its accuracy, stability and efficiency. Unlike Smoothed Particle Hydrodynamics (SPH), which purely relies on particles to discretize the computational domain, FLIP/APIC uses both particles and a background Eulerian grid. The Navier-Stokes equations are solved on the grid, allowing for accurate derivative calculations, well-defined free surface and solid boundary conditions, and accurate first-order approximation of physical reality. The FLIP/APIC method also circumvents common artifacts of SPH; e.g., underestimated density near free surfaces and weakly compressible artifacts. In fact, the requirement for incompressibility is crucial in the fluid-pouring problem studied in this paper. We choose not to use SPH because it does not guarantee a divergence-free velocity field unless additional computational components are included. FLIP/APIC, however, 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 and the tracking of material properties. Besides modeling fluid, the state-of-the-art physics-based simulation methods have provided realistic cues for modeling complex tool and tool-uses (Zhu, Zhao, & Zhu, 2015), generic containers (Liang, Zhao, Zhu, & Zhu, 2015) and soft human body (Zhu, Jiang, Zhao, Terzopoulos, & Zhu, 2016).

Realistic visualization of simulated fluids is particularly important for the flow demonstration animations displayed before each trial in our viscous fluid-pouring task. To provide vivid impressions of the motion of the *LV* and *HV* fluids, we adopt OpenVDB (Museth et al., 2013) and utilize the latest particle fluid surfacing methods developed in the field of computer graphics to reconstruct a smooth level set surface from the simulated fluid particles based on their locations.

The favorable efficiency and precision of the FLIP/APIC method allows for effortless generation of ground truth responses for our task, given that fluid viscosity and volume are known. The particle locations and vessel tilt angles are explicitly recorded at each time step as simulation outputs. Since the FLIP/APIC method does not involve any stochastic processes, the output of each simulation is deterministic. In each simulation, 40,000 particles (with around 8 particles per grid cell) were used to ensure both stability and convergence.

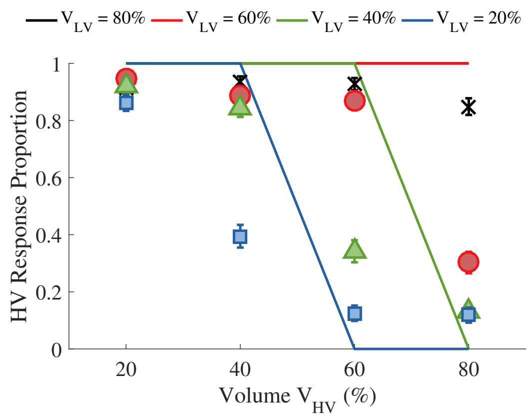

Figure 3: Simulation results from the FLIP/APIC model ($RMSD = 0.6747$). Separate lines indicate model predictions, given ground-truth volume/viscosity values for each fluid. Symbols indicate human response proportions. The deterministic simulation method returns binary predictions and does not provide probabilistic response proportion estimates.

Intuitive Fluid Engine

Fluid simulation with physical dynamics provides deterministic fluid movements if the ground-truth values of viscosity and volume are known. Hence, the decisions directly derived from the FLIP/APIC fluid simulator are binary judgments (Fig. 3), which implies that the physical simulation with high precision cannot explain humans' probabilistic judgment in the fluid-pouring task. Inspired by the approach of Bates et al. (2015) and the noisy Newton model (e.g., Sanborn et al., 2013), we combine the physical simulator of FLIP/APIC with noisy input variables (i.e., viscosity and volume) to form the Intuitive Fluid Engine (IFE) model, thereby accounting for physical uncertainty and the influence of viscosity and volume on our reasoning task.

The input variables for our IFE are the ground-truth values of volume and viscosity (V_T, μ_T) for the two fluids in each experimental trial. The sampling process of the IFE then samples $N = 10,000$ noisy viscosity and volume pair values $\{(V_i, \mu_i), i = 1, 2, \cdots, N\}$ for each fluid. Each sample (V_i, μ_i) of the 10000 generated noisy input variables is later passed to the FLIP/APIC simulator to produce a binary decision $B_i(V_i, \mu_i) \in \{L, R\}$. The decision is that either the left L or the right R container needs to be tilted with a larger angle before the fluid inside begins to pour out. By aggregating the prediction from all $10,000$ samples and dividing the sum by the number of samples, the IFE outputs the distribution $\boldsymbol{P}(V_T, \mu_T)$ for the given ground-truth values of viscosity and volume:

$$\boldsymbol{P}(V_T, \mu_T) = \begin{cases} P(V_T, \mu_T)_L = \dfrac{\sum_i H(B_i(V_i, \mu_i), L)}{N} \\ \\ P(V_T, \mu_T)_R = \dfrac{\sum_i H(B_i(V_i, \mu_i), R)}{N}, \end{cases} \tag{3}$$

where $H(\Psi, \Theta) = 1$ when $\Psi = \Theta$, and it is 0 otherwise.

To model physical uncertainty, the sampling process of the IFE model is implemented by adding perceptual noises to the ground-truth values of the physical input variables (i.e., viscosity and volume). Noisy volume is generated by adding an offset to its ground-truth value from a Gaussian distribution with mean 0 and variance σ_μ, whereas the noisy viscosity is generated by adding a fixed amount of Gaussian noise on a logarithmic scale (Sanborn et al., 2013): $V_i = f^{-1}(f(V_T) + \varepsilon)$, where V_T is the ground-truth value, $f(V_T) = \log(\omega \cdot V_T + k)$, f^{-1} is the inverse of f, and $\varepsilon \sim \text{Gaussian}(0, \sigma_V)$. The results reported herein chooses $\sigma_\mu = 0.1$, $\sigma_V = 0.1$, $k = 1.5$, $\omega = 1$.

Each simulation required approximately 10 minutes to run on a modern single-core CPU. In order to run a large number of simulations during the sampling process, we discretized the viscosity and volume spaces into finite sets. Specifically, the viscosities are discretized into 8 different cases $(0.1, 1, 10, 100, 200, 500, 1000, 2000)$ and the volumes into 21 different cases (0% to 100% with a step-size 5%). We precomputed the simulation results for each discretized case, and stored the results in a database. During the sampling process, the sampled inputs are numerically rounded to the precomputed discretized cases, where the results can be immediately retried from the database without re-computing.

Non-Simulation Models

To examine whether fluid 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) (McCullagh, 1984) and the support vector machine (SVM) (Cortes & Vapnik, 1995). These models are purely data-driven and do not involve any explicit knowledge of physical laws or physical simulation. 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.

To predict human judgment for the $i^\star$th trial $J_{i^\star}$, both non-simulation models were tested with the $i^\star$th trial, and trained with the remaining 15 trials $\{J_i, i = 1, 2, \cdots, 16, i \neq i^\star\}$. The trained GLM model is directly applied to the test case to predict which container will need to be tilted to a larger angle before the fluid inside begins to pour out. Since the SVM is a discriminative classification method which can only predict discretized labels (i.e., $+1$ indicating selection of the left container and -1 indicating selection of the right container), we introduced perceptual noise (the same method for the IFE) to each test trial to model physical uncertainty. For each test trial, a set with $10,000$ samples was generated. The trained SVM model is then applied to predict the labels ($+1$ or -1) in each sample, which are then aggregated to form the probability distribution for each test trial.

Model Results

We first compared how well different computational models account for human performance for the 16 trials. Fig. 4 depicts results from the IFE, GLM, and SVM models with perceptual noise. Human judgments and model predic-

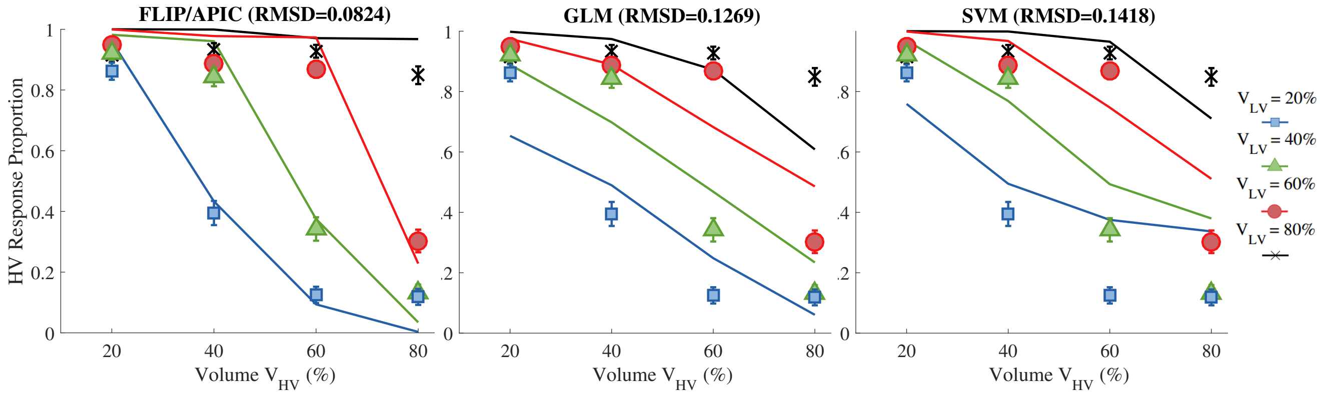

Figure 4: Comparison of results between our three prediction models: (Left) IFE, (Middle) Regression, and (Right) SVM with perceptual noise. Horizontal axes denote *HV* fluid volume; vertical axes denote the predicted proportion of *HV* fluid responses associated with a greater rotation angle. The IFE simulation model outperforms competing data-driven models.

tions were highly correlated ($r(14) = 0.9954, 0.9488$, and 0.9251, respectively). RMSD (root-mean-squared deviation) between human judgments and the models' predictions are $0.0824, 0.1269$, and 0.1418, respectively. Compared to the purely data-driven models (i.e., the GLM and SVM models), the simulation-based IFE model encodes material properties (e.g., viscosity) and perceptual features (e.g., volume) and provides better approximations to human judgments in the viscous fluid-pouring task. These results again support the role of simulation as a potential mental model that supports human inference in physical reasoning tasks.

Next, we examined whether the IFE model captures human performance on the three trials where the heuristic rule outlined earlier provides incorrect predictions; i.e., those trials where V_{HV} = 40%, 60%, and 80% and V_{LV} = 20%, 40%, and 60%, respectively. Here, the ground-truth model predicts that the *HV* fluid will require a greater angle of rotation before beginning to pour, while the heuristic rule discussed earlier suggests the opposite. *HV* response proportions for these trials were .39, .34, and .30, respectively, and the IFE model returned consistent predictions of .43, .37, and .23. Alternatively, the GLM model predicted response proportions of .49, .47, and .49, and the SVM model predicted response proportions of .49, .49, and .51. Thus, our IFE model captured human performance on the specified trials while competing data-driven methods returned predictions biased toward the ground-truth model and away from the lesser-volume heuristic.

Discussion

Our results from the viscous fluid-pouring task agree with the findings of Bates et al. (2015) in that our probabilistic, simulation-based IFE model outperformed two non-simulation models (SVM and GLM). Our behavioral experiment also indicates that people naturally attend to latent attributes (e.g., viscosity) when reasoning about fluid states following observation of realistic flow demonstration animations. By extending our probabilistic IFE method to a reasoning task, we demonstrated that the noisy Newton framework can account for human performance in a fluid-related judgment task that traditionally precludes mental simulation strategies.

While simulation has been demonstrated as the default strategy in other mechanical reasoning tasks (Hegarty, 2004; Clement, 1994), the participants in Schwartz and Black's (1999) experiments failed to spontaneously represent and simulate physical properties relevant to the water-pouring problem when making their judgments. It is important to note, however, that the present task differs from the traditional water-pouring task in several ways: (i) fluid viscosity and volume (rather than container diameter) varied across trials, (ii) a cup-tilting demonstration was displayed to visualize the rate of simulated rotation, and (iii) motion cues from flow demonstrations informed the perception of latent fluid attributes (e.g., viscosity). Comparison of our study to previous water-pouring studies suggests that the dissociation between explicit physical prediction and implicit judgment reported in the intuitive physics literature could be resolved in some situations by modifying task characteristics and instructions in ways that motivate simulated representation. While our viscous water-pouring problem indicates a set of simulation-inducing task characteristics, further research should aim to determine specific experimental factors that trigger simulation strategies. Specifically, can the conditions employed in the present task extend to classical rigid-body and fluid mechanics problems to resolve the discrepancy between people's explicit predictions and tacit judgments, and if so, what additional task characteristics serve to facilitate mental simulation?

Classical research in artificial intelligence has traditionally dismissed robust mental simulation as a strategy for physical reasoning due to its inherent complexity, often proposing simplified qualitative models instead (De Kleer & Brown, 1984). While the computational fluid simulations employed in the present study require extensive numerical evaluation to make predictions about future fluid states, humans appear to do so with precision and accuracy in comparatively small amounts of time. Furthermore, their performance in our reasoning task suggests representation of physical quantities that extends beyond qualitative process theory. 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 (Grzeszczuk, Terzopoulos, & Hinton, 1998).

Acknowledgments

Support for the present study was provided by an NSF Graduate Research Fellowship, DoD CASIT grant W81XWH-15-1-0147, DARPA SIMPLEX grant N66001-15-C-4035 and ONR MURI grant N00014-16-1-2007.

References

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

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

Bridson, R. (2008). *Fluid simulation for computer graphics*. CRC Press.

Clement, J. (1994). Use of physical intuition and imagistic simulation in expert problem solving.

Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine learning*, *20*(3), 273–297.

De Kleer, J., & Brown, J. S. (1984). A qualitative physics based on confluences. *Artificial Intelligence*, *24*(1), 7-83.

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

Grzeszczuk, R., Terzopoulos, D., & 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.*

Hegarty, M. (2004). Mechanical reasoning by mental simulation. *Trends in cognitive sciences*, *8*(6), 280–285.

Howard, I. P. (1978). Recognition and knowledge of the water-level principle. *Perception*, *7*(2), 151.

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

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

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

Krist, H., Fieberg, E. L., & Wilkening, F. (1993). Intuitive physics in action and judgment: The development of knowledge about projectile motion. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*(4), 952.

Liang, W., Zhao, Y., Zhu, Y., & Zhu, S.-C. (2015). Evaluating human cognition of containing relations with physical simulation. In *Proceedings of the 37th annual conference of the cognitive science society.*

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

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

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

Museth, K., Lait, J., Johanson, J., Budsberg, J., Henderson, R., Alden, M., ... Pearce, A. (2013). Openvdb: an open-source data structure and toolkit for high-resolution volumes. In *Acm siggraph 2013 courses* (p. 19).

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

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

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

Smith, K., Battaglia, P., & Vul, E. (2013). Consistent physics underlying ballistic motion prediction. In *Proceedings of the 35th conference of the cognitive science society* (pp. 3426–3431).

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

Tversky, B., Morrison, J. B., & Betrancourt, M. (2002). Animation: can it facilitate? *International journal of human-computer studies*, *57*(4), 247–262.

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

Zhu, Y., Jiang, C., Zhao, Y., Terzopoulos, D., & Zhu, S.-C. (2016). Inferring forces and learning human utilities from videos. In *Ieee conference on computer vision and pattern recognition (cvpr).*

Zhu, Y., Zhao, Y., & Zhu, S.-C. (2015). Understanding tools: Task-oriented object modeling, learning and recognition. In *Ieee conference on computer vision and pattern recognition (cvpr).*