
Learning Physics by Asking: Revisable Beliefs for Reconstructed 3D Assets

Qinzheng Ma
Rice University
Houston, TX, USA
qm18@rice.edu

Abstract

When is a reconstructed object’s physical description sufficient for a robot task? We propose treating each asset as a persistent belief over physical parameters, with measurement provenance and a record of which tasks have been independently tested. For rigid tabletop objects with known geometry we study mass and contact-pair kinetic friction: visual hypotheses initialize the belief, task-conditioned probes revise it, and unidentifiable directions are retained. An exact four-hypothesis analysis shows that the probe with more parameter information can have zero value for the current task, and that successful sliding cannot establish readiness for lifting. A numerical pilot in a hidden-parameter reference environment (200 synthetic objects, no robot data) tests the resulting predictions: under a light-biased visual prior, five slides leave the mass interval uncalibrated while one task-selected lift raises its coverage from 0.62 to 0.83, task-risk selection reaches near-oracle lifting regret with about one contact where parameter information gain spends its first contact on friction, and a passed lift task can leave a $2.5\times$ mass error that a later push task exposes. A language agent orchestrating the same tools is a matched-control research question, not a premise.

1 Introduction and Related Work

A visually plausible asset can predict a slide correctly while assigning the wrong mass. Successful replay does not resolve this ambiguity, and a later lifting task may expose it. We ask how limited interaction evidence should change an object’s physical hypotheses, which uncertainty must persist, and what supports a task-specific claim of adequacy. Our thesis is that *physical completion is a revisable state of knowledge, not a terminal assignment of parameters*.

PhysX-Anything generates physical assets from an image [3]; Scalable Real2Sim identifies geometry and inertial properties through robot interaction [8]. VLM-generated behavior trees already acquire parameters according to task intent [1]. Agentic Real2Sim preserves observations and evaluation traces while constructing episodic twins [5]; provenance alone therefore cannot establish novelty. ASID connects informative exploration, system identification, and policy learning [6]. Bayesian design [4] and dynamics randomization [7] are standard.

We propose an explicit *knowledge lifecycle*: preserve evidence-equivalent hypotheses, reopen task adequacy after a task or surface change, and account for the cost of both identification and independent validation. Our contributions are (i) a concrete lifecycle specification, (ii) an exact diagnostic separating parameter information from decision value, (iii) a numerical pilot that tests the diagnostic’s predictions about revision, evidence cost, and transfer across tasks, and (iv) an evaluation protocol that isolates agent orchestration and uncertainty propagation; none is a robot result or a priority claim.

The developmental motivation concerns changing expectations with experience. PLATO links object-level representations to physical prediction [9]; studies of confounded evidence [10] and prior-dependent exploration [2] motivate distinguishing hypotheses and testing counterevidence. The behavioral signature we test is developmental in form, not mechanism: an expectation formed from appearance, an intervention that cannot discriminate the relevant hypotheses, retention of that ambiguity in the object record, and revision by an intervention that can; we do not model children’s cognitive mechanisms or claim new physical concepts.

2 Proposed Framework

Objects, beliefs, and evidence. Inputs are a metric mesh and collision model G , images V , a task T , calibrated tool metadata, and an exploration budget B . The first study covers nonrotating tabletop translation and stable vertical lifting, with unknowns $\theta = (m, \mu_k)$ for a named object–surface pair. The object record holds a belief $b_t(\theta) = p(\theta \mid G, V, D_t)$ and an evidence ledger D_t . VLM-suggested ranges map to a distribution uniform in log-mass and friction, mixed with a broader bounded fall-back, with mapping and weight fixed on development objects. Each measurement stores identifiers, controls, units, raw data, calibration and likelihood versions, and source, so calibration changes force replay and geometry changes mark validations stale. Primitives execute a free slide with measured initial velocity or obtain a tared quasi-static lifting force. A Gaussian summary likelihood $p(y \mid \theta, e)$ combines a physics prediction with sensor noise and a development-fixed discrepancy allowance; its mixture $p_b(y \mid e) = \int p(y \mid \theta, e)b(\theta)d\theta$ predicts possible evidence. Only real sensors, external measurements, or a learner-independent hidden-parameter reference environment can supply D_t ; candidate simulations are predictions, never self-generated evidence.

Selection, inference, and stopping. Task T defines an initial-condition distribution $q_T(\xi)$, a fixed controller bank Π_T , and a loss $\ell_T \in [0, 1]$ that includes constraint failure and, when declared, a delay penalty. We use standard myopic decision value:

$$R_T(b) = \min_{\pi \in \Pi_T} \mathbb{E}_{\theta \sim b, \xi \sim q_T} [\ell_T(\pi, \theta, \xi)], \quad (1)$$

$$U_T(e; b) = \frac{R_T(b) - \mathbb{E}_{y \sim p_b(\cdot \mid e)} R_T(b^{e,y})}{c(e)}. \quad (2)$$

$b^{e,y}$ is a hypothetical posterior; $c(e) > 0$ weights contacts, resets, and compute. A numerical module samples hypothetical observations, reweights a log-mass/friction grid that retains likelihood-equivalent directions, and recomputes the finite minimum. The agent selects a task template, requests feasible probes and scores, takes the best feasible probe, and replans after tool failures. With identical candidates and scores, language reasoning cannot improve on an exhaustive numerical maximizer of Eq. (2), so agent value must be tested in tool orchestration, including its overhead. An independent probe updates $b_{t+1}(\theta) \propto p(y_t \mid \theta, e_t)b_t(\theta)$; before normalization a standardized-residual test checks whether *any* supported model fits, and if none does the system flags mismatch, checks calibration and geometry, and either re-infers under a versioned support expansion or abstains. Enlarging noise until a fit appears is not validation.

Stopping (budget exhaustion, tool unavailability, or a fixed utility threshold) is a resource decision. A separate, fixed validation allocation tests prediction error and task failure under q_T with frozen controller and belief, and never chooses the stopping point or fits the belief. The record distinguishes *unverified*, *passed within the tested scope*, and *model-inadequate*; a pass requires predeclared error tolerances and a failure upper bound below the task threshold, which posterior concentration cannot replace. Export includes geometry, joint weights, lineage, unresolved directions, and test scope; training samples one joint $\theta \sim b_t$ per episode. A new task reuses the belief but reopens validation; a new surface retains mass and starts a new contact-pair friction belief.

3 Exact Analytical Diagnostic

An identification limit. Assume a horizontal rigid body, no rotation or air resistance, constant Coulomb kinetic friction, known initial speed $v_0 > 0$, and normal force mg . Until stopping,

$$m\ddot{x} = -\mu_k mg, \quad x(t) = x_0 + v_0 t - \frac{1}{2}\mu_k g t^2, \quad d = \frac{v_0^2}{2\mu_k g}. \quad (3)$$

With mass-independent measurement noise, masses at the same μ_k have identical slide likelihoods, and by induction this holds for any adaptive finite sequence of slides, since equal history likelihoods induce the same next probe and each observation remains mass-independent. A factorized prior retains its mass marginal; a correlated prior may shift it without identifying mass. A stable lift clear of the table instead measures $F_z = m(g + a_z)$, or mg quasi-statically, after taring.

Information and decision value can disagree. Consider four hypothetical states: $m \in \{0.20, 0.40\}$ kg and $\mu_k \in \{0.15, 0.30\}$, with independent priors $P(m = 0.40) = p$ and uniform friction, $g = 9.81 \text{ m/s}^2$. For $\mu_k = 0.30$ and $v_0 = 0.60 \text{ m/s}$ both masses stop at $d_* = 0.061162 \dots \text{ m}$. Sliding task S chooses $v \in \{0.60/\sqrt{2}, 0.60\} \text{ m/s}$ to stop within 1 cm of d_* with binary failure loss; each speed succeeds at exactly one friction, so $R_S = 1/2$. Lifting task L chooses a rest-to-rest profile over 0.10 m with acceleration $a \in \{2, 8\} \text{ m/s}^2$ under a 6 N peak-force bound; the slow profile is feasible for both masses, the fast one only for the light mass, and the loss is 1 for infeasibility, otherwise 0 for fast and 1/4 for slow. Without this delay preference, always-slow would be optimal and mass unnecessary. For ideal noiseless unit-cost probes, a slide reveals friction and a quasi-static lift reveals mass, and direct enumeration gives $R_L = \min(p, 1/4)$, $\mathbb{E}_y R_L(b^{\text{lift}, y}) = p/4$, and $U_L(\text{lift}) = \min(p, 1/4) - p/4$. At $p = 1/20$ the slide carries 1 bit of parameter information with $U_S = 1/2$ and $U_L = 0$, whereas the lift carries $h_2(1/20) \approx 0.286$ bits with $U_S = 0$ and $U_L = 3/80$. Information gain therefore prefers sliding even for task L , where it cannot improve the decision, while mass is unnecessary for task S ; a task switch changes decision relevance, not what prior slides established. With noisy lifting, posterior mass odds equal prior odds times the likelihood ratio: at $p = 1/20$, making heavy more probable than light requires a ratio above 19, so one noisy lift need not overturn a confident prior.

4 Numerical Pilot in a Hidden-Parameter Reference Environment

We test the diagnostic’s predictions (slides never revise mass; information gain can spend a budget on the wrong probe; task-risk selection reaches the task standard at lower cost) in a reference environment that holds hidden (m^*, μ^*) per object, m^* log-uniform on $[0.1, 1] \text{ kg}$ and μ^* uniform on $[0.1, 0.6]$, and returns stopping distances (noise sd 4 mm) for slides launched at 0.3, 0.5, or 0.7 m/s and tared quasi-static forces (sd 0.2 N) for lifts; every probe costs one contact. The learner keeps a 201×161 log-mass/friction grid with the same mechanics and noise (a well-specified likelihood). The visual mass prior is log-normal (sd 0.25) mixed 0.85:0.15 with a bounded uniform fallback; the *correct* condition centers it at m^*e^ϵ , $\epsilon \sim \mathcal{N}(0, 0.2^2)$, and the *misleading* condition at $m^*/2.5$, a confident light-biased guess. Friction priors are Gaussian (sd 0.12) around a noisy μ^* (sd 0.1), mixed 0.8:0.2 with uniform. Task S must stop within 1.5 cm of a target $\xi \sim \mathcal{U}[0.10, 0.20] \text{ m}$ using a bank of assumed frictions; task L lifts rest-to-rest at 2, 4, or 8 m/s² under a 10 N peak-force bound with delay losses 1/4, 1/8, 0; task P pushes with a force from a 0.1–9 N bank and fails unless the net acceleration $F/m - \mu_k g$ lies in $[0.5, 2] \text{ m/s}^2$, so it needs both parameters. Seven selectors share priors, tools, and common observation noise: slide-only, lift-only, and alternating fixed protocols, random probes, parameter information gain (512 hypothetical observations per candidate), and myopic task-risk per unit cost (Eq. (2), same samples) with and without a stopping rule (best utility below 10^{-3}). After each budget $b \leq 5$ we freeze the belief, pick the risk-minimizing controller, and report regret against the oracle controller that knows θ^* , 90% mass-interval coverage, and contacts used over 200 objects. A task-switch experiment then runs $S \rightarrow L \rightarrow P$ with the stopping selector, either carrying the record across tasks (*persistent*) or re-initializing it from the visual prior at each switch (*reset*). No language agent runs.

Findings. *Slides never revise mass.* Under the misleading prior, five slides leave mass coverage at 0.62 and lifting regret at 0.184, their prior-only values, while reducing sliding regret from 0.731 to 0.008: the evidence that completes task S leaves task L where it started. One lift raises coverage to 0.83–0.84 and cuts lifting regret to 0.006; two lifts give 0.89. *Information gain spends contacts on the wrong probe.* For the lift task it chooses a slide first for 88% of objects, so regret after one contact is 0.135 against 0.019 for task-risk selection; for the slide task it spends 2.6 of 5 contacts on lifts and ends at 0.037 against 0.008. *Task-risk selection matches the best fixed protocol for each task without being told which it is.* Lift-only on L and slide-only on S are the ceilings; task-risk without stopping tracks both (0.006 and 0.008 after five contacts), and the stopping rule trades a little regret for cost (1.16 contacts on L and 3.78 on S at 0.018 and 0.008). A second seed and the

Selector	Lift task L					Slide task S			
	r_1	r_3	r_5	cov_5	c_5	r_1	r_3	r_5	c_5
Slide only (fixed)	0.184	0.184	0.184	0.62	5.00	0.096	0.020	0.008	5.00
Lift only (fixed)	0.006	0.006	0.006	0.89	5.00	0.731	0.731	0.731	5.00
Alternate slide/lift (fixed)	0.184	0.003	0.005	0.89	5.00	0.096	0.038	0.024	5.00
Random probe	0.142	0.095	0.078	0.78	5.00	0.431	0.202	0.118	5.00
Parameter information gain	0.135	0.003	0.009	0.87	5.00	0.160	0.076	0.037	5.00
Task-risk per cost, no stopping	0.011	0.006	0.006	0.88	5.00	0.096	0.020	0.008	5.00
Task-risk per cost, stops (Eq. 2)	0.019	0.018	0.018	0.83	1.16	0.096	0.020	0.008	3.78

Table 1: Misleading prior, $n = 200$ hidden objects per cell. r_b : mean regret versus the oracle controller after budget b (prior-only 0.184 for L , 0.731 for S ; s.e. ≤ 0.033); cov_5 : 90% mass-interval coverage after five contacts (prior-only 0.62); c_5 : contacts used. Reference-environment values, not robot data.

correct prior preserve the ordering. *Persistence pays only where evidence overlaps, and a pass is not identification.* In $S \rightarrow L \rightarrow P$, carrying the record gives L no head start over resetting (slides carry no mass evidence), but leaves P at regret 0.185 before any push-phase contact against 0.990 after a reset, and 2.20 against 4.00 contacts to finish (7.21 against 8.97 in total; 7.12 against 8.92 with the correct prior). The residual 0.185 is a passed task reopened: the 2% of objects whose lift profile the prior had already settled passed L without a lift and start P at regret 1.000, and light objects whose single 0.2 N-noise lift fixed the profile but not the mass precision P needs start at 0.164; P weighs both again. Two cautions: after one lift under the confident misleading prior, coverage is 0.83 rather than the nominal 0.90 because light objects stay pulled toward the visual guess (the noisy-lift caveat of Section 3); and with 128 rather than 512 hypothetical samples, Monte Carlo error sometimes stopped the task-risk selector before any probe, so myopic stopping under a wrong prior is a hazard, never evidence of adequacy.

5 Remaining Protocol and Hypotheses

With zero failures in n independent trials, the exact one-sided 95% binomial bound on the failure rate is $1 - 0.05^{1/n}$: five successes only bound it below 0.451, and a 5% bound needs 59 trials, a cost that applies to every row of Table 1; low regret after one contact is not a validated pass.

H1: Evidence revision corrects misleading visual priors without erasing ambiguity. Supported by the pilot, including its task-switch and reset conditions, for a well-specified model; remaining tests add force-sensor removal and real or independently simulated assets, reporting physical-unit errors, coverage, and failure-event Brier scores.

H2: Task-conditioned acquisition reduces total evidence cost at a fixed task standard. The pilot compares numerical selectors only; the open comparison adds agent orchestration with identical priors, inference, tools, and budgets, reporting contacts, resets, compute, latency, non-attainment, and false passes on final independent tests; equal or worse cost than the task-risk comparator undermines the agent, and gains bought with extra sensing or compute do not count.

H3: Posterior sampling improves training under held-out dynamics. Untested. For identical acquired evidence, compare posterior sampling, a point estimate, and broad-prior randomization under the same algorithm, budget, and at least three seeds, with reference parameters hidden and test trajectories held out; equal or better point-estimate performance undermines H3.

Limitations. The pilot shares the learner’s model with the environment, omits rotation, geometry error, and contact variation, and runs no language agent: it shows that revision and task-conditioned selection behave as predicted, not that the full system works on a robot. Known equations and a small controller bank constrain learning, so the agent may prove unnecessary. A passed test does not identify every parameter or validate another controller, and posterior training may inherit model error. Our developmental scope is persistent expectations, intervention, revision, and transfer within a fixed physical model; real-world transfer and human-like understanding remain unestablished. For developmental research the pilot is a scored testbed for two behaviors that studies of children’s exploration describe qualitatively: retaining an ambiguity the evidence cannot resolve, and seeking discriminating evidence only when a new task demands it.

References

- [1] Alessandro Adami, Sebastian Zudaire, Ruggero Carli, and Pietro Falco. Real2Sim via active perception with behavior trees automatically generated by VLMs. arXiv:2601.08454v2, 2026. URL <https://arxiv.org/abs/2601.08454v2>.
- [2] Elizabeth Baraff Bonawitz, Tessa J. P. van Schijndel, Daniel Friel, and Laura E. Schulz. Children balance theories and evidence in exploration, explanation, and learning. *Cognitive Psychology*, 64(4):215–234, 2012. doi: 10.1016/j.cogpsych.2011.12.002. URL <https://eccl.scripts.mit.edu/papers/BonawitzVanSchijndelFrielSchulz2012.pdf>.
- [3] Ziang Cao, Fangzhou Hong, Zhaoxi Chen, Liang Pan, and Ziwei Liu. PhysX-Anything: Simulation-ready physical 3D assets from single image. arXiv:2511.13648v1, 2025. URL <https://arxiv.org/abs/2511.13648v1>.
- [4] Kathryn Chaloner and Isabella Verdinelli. Bayesian experimental design: A review. *Statistical Science*, 10(3):273–304, 1995. doi: 10.1214/ss/1177009939. URL <https://doi.org/10.1214/ss/1177009939>.
- [5] Guanxiong Chen, Qianjun Xia, Jiawei Peng, Heng Zhang, Bole Ma, Justin Qian, Ziyi Jiao, Bingyang Zhou, Luoxin Ye, Kaifeng Zhang, Kunyi Wang, Weijia Zeng, Yunuo Chen, Pengzhi Yang, Ziqiu Zeng, Siyuan Luo, Huamin Wang, Chao Liu, Alan Yuille, Fan Shi, Changxi Zheng, Yunzhu Li, Chenfanfu Jiang, and Peter Yichen Chen. Agentic Real2Sim: Physics-based world modeling with vision-language agents. arXiv:2607.19190v3, 2026. URL <https://arxiv.org/abs/2607.19190v3>.
- [6] Marius Memmel, Andrew Wagenmaker, Chuning Zhu, Patrick Yin, Dieter Fox, and Abhishek Gupta. ASID: Active exploration for system identification in robotic manipulation. arXiv:2404.12308v2, 2024. URL <https://arxiv.org/abs/2404.12308v2>.
- [7] Xue Bin Peng, Marcin Andrychowicz, Wojciech Zaremba, and Pieter Abbeel. Sim-to-real transfer of robotic control with dynamics randomization. arXiv:1710.06537v3, 2018. URL <https://arxiv.org/abs/1710.06537v3>.
- [8] Nicholas Pfaff, Evelyn Fu, Jeremy Binagia, Phillip Isola, and Russ Tedrake. Scalable Real2Sim: Physics-aware asset generation via robotic pick-and-place setups. arXiv:2503.00370v2, 2025. URL <https://arxiv.org/abs/2503.00370v2>.
- [9] Luis S. Piloto, Ari Weinstein, Peter Battaglia, and Matthew Botvinick. Intuitive physics learning in a deep-learning model inspired by developmental psychology. *Nature Human Behaviour*, 6:1257–1267, 2022. doi: 10.1038/s41562-022-01394-8. URL <https://www.nature.com/articles/s41562-022-01394-8>.
- [10] Laura E. Schulz and Elizabeth Baraff Bonawitz. Serious fun: Preschoolers engage in more exploratory play when evidence is confounded. *Developmental Psychology*, 43(4):1045–1050, 2007. doi: 10.1037/0012-1649.43.4.1045. URL <https://eccl.scripts.mit.edu/papers/bonawitzandschulzseriousfun.pdf>.