The Patchwork Approach: Toward a Perceptual Theory of Intuitive Physics

Abdul-Rahim Deeb
Department of Cognitive Science
Johns Hopkins University

Abstract

Human experience is rich with sensory information, from which physical regularities are internalized to guide behavior. This paper presents the Patchwork Approach, a method for modeling sensorimotor predictions based on these regularities, without explicitly encoding physics. This method leverages environmental regularities, enabling resource-rational perceptual predictions. Using data from previous studies (Deeb et al., 2021, 2024) on human perception, the model outperforms Newtonian-based models in capturing human prediction errors and interpolates across 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 from untrained aiming angles. This approach provides valuable insights into how physical laws are internalized and used to guide perception and action–arguing 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).

Keywords: intuitive physics; perception; sensorimotor prediction; computational modeling

Introduction

Learning through experience is crucial for successful sensorimotor interactions with the physical world. Whether it's a little leaguer learning to pitch or a server stacking dishes, repeated experiences refine one's ability to perceive and adjust to physical interactions. The deterministic nature of physical systems raises the question of whether humans leverage these regularities, identify anomalies, and represent or *internalize* the correlations needed to form accurate predictions about the physical world. Repeated exposure to the physical world may implicitly represent key principles of motion, force, and spatial relationships, enabling better predictions.

Research on this topic has traditionally focused on whether explicit human reasoning aligns with classical mechanics, revealing significant discrepancies between human judgment and established physical laws (McCloskey et al., 1980). Recent approaches, such as the Intuitive Physics Engine (IPE), have attempted to reconcile human reasoning with physical ground truth (Battaglia et al., 2013) by using a mental model of the physical world that simulates future events, leveraging probabilistic reasoning and robust generalization to capture human judgment (Wang et al., 2024).

Often overlooked is the considerable precision and accuracy with which humans perceive and act upon physical systems in daily life. Previous research has demonstrated that, unlike higher-order reasoning, our perceptual system is finely tuned to the physical laws governing interactions in our environment. For instance, individuals excel in tasks involving projectile motion (Kaiser et al., 1985; McIntyre et al., 2001; Monache et al., 2019; Montagne et al., 1999; Siegler et al., 2010), recovering object masses (Deeb & Domini, 2025; Mitko & Fischer, 2023; Mitko et al., 2024; Ross, 1969; Runeson & Frykholm, 1983; Sanborn et al., 2013; Todd & Warren, 1982), and anticipating the behavior of launched objects (Badler et al., 2010; Deeb et al., 2022; Kim et al., 2013; Michotte, 1963; Rolfs et al., 2013). Moreover, these sensorimotor capacities likely tap into different mechanisms than those responsible for explicit reasoning (Fischer et al., 2016; Kaiser et al., 1992; Krist et al., 1993; Mason & Just, 2016).

Recent work in motion perception has suggested that physical predictions themselves act as cues that can be integrated into sensory information (Deeb et al., 2021; Deeb & Domini, 2024), framing perception as a foundational input to intuitive physics reasoning—complementing higher-level cognitive processes rather than serving as a subordinate precursor. Despite this growing body of evidence that physical laws are embedded within sensory processes that underlie perceptual experience, formalizing a robust perceptual theory of intuitive physics has proven challenging. Recently, the PLATO model has emerged as a powerful approach to explaining how physical expectations are developed through Violation-of-Expectation paradigms, capturing how discrepancies between predicted and observed outcomes inform learning (Piloto et al., 2022). However, more work is needed to explore how specific quantitative predictions and latent physical features are embedded within sensory processes that underlie perceptual experience.

The Patchwork Approach and Physical Representations

The Patchwork Approach emphasizes the identification of relationships between perceivable events, rather than strict adherence to formal definitions of physical laws.This approach is fundamentally concerned with how individuals leverage correlations in the environment to generate predictions about physical events. By offloading the representation of physical contingencies onto the environment—which operates according to established physical principles—individuals can intuitively navigate their surroundings based on experienced interactions, without needing a detailed understanding of the

In D. Barner, N.R. Bramley, A. Ruggeri and C.M. Walker (Eds.), *Proceedings of the 47th Annual Conference of the Cognitive Science Society*

underlying laws. While classical mechanics provides a precise framework for understanding physical phenomena, the statistics of human experience reflect correlations that are not always properly defined by these formal laws. As a result, the correlations we internalize—and use to guide our predictions—may diverge from the principles of classical mechanics, yet remain adaptive within the contexts in which they are learned. This is exemplified by the mass-speed illusion (Deeb & Fischer, 2025), which shows how perceived object mass influences velocity judgments, highlighting how human experience reveals correlations beyond classical mechanics—such as the perception that heavier objects tend to move slower or that smaller objects are denser (Peters et al., 2015) or that darker objects are heavier (Walker et al., 2010).

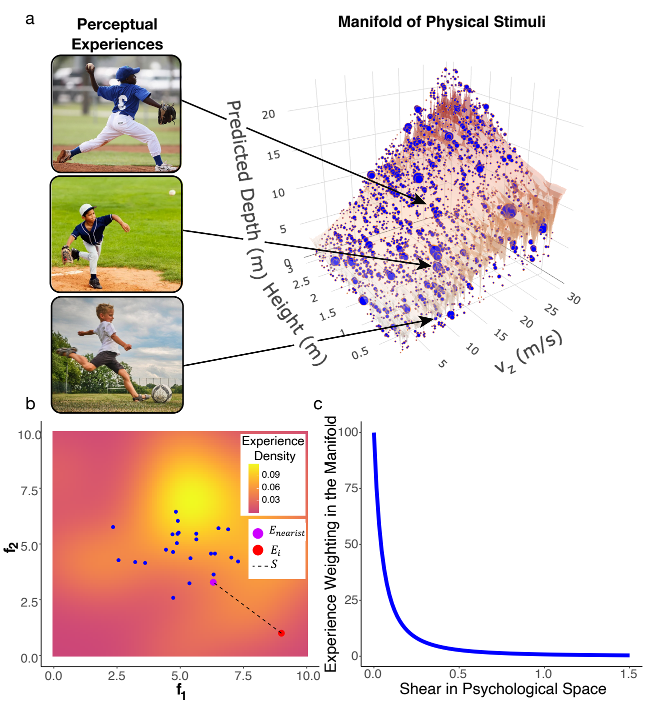

Figure 1: a) The Patchwork Approach framework: Experiences are integrated into a multi-dimensional manifold, with behaviorally relevant features (e.g., time-to-contact, location in depth, initial height, velocity) forming the dimensions. The right plot shows how experiences (depicted as blue points) are arranged. (b) Distribution of experiences in a generic feature space, with the closest experience (magenta) influencing the new experience (red). The dashed line represents the distance between the new experience and its nearest neighbor. (c) The weighting function decreases as the distance from the nearest experience increases, with a sharp decline corresponding to greater psychological shear.

Building on Shepard's (1987; 1994) proposal that psychological space is structured by universal principles reflecting the regularities of the physical world, the Patchwork Approach offers a novel perspective on how humans internalize and apply these principles in perception and action. Physical perceptual representations are conceptualized as a manifold with multiple dimensions, where each dimension reflects behaviorally relevant, observable features—such as speed, size, and time—shaped by a range of experiences, and may vary across developmental stages and contexts (Bremner et al., 2012). By leveraging correlations among these features, individuals can form flexible, context-sensitive laws that guide their understanding of object behavior, enabling physical predictions that integrate into perceptual experiences. Like the PLATO model, the Patchwork Approach relies on structured representations that organize experiences in ways responsive to the natural contingencies of the world. However, it uniquely emphasizes the role of contingencies between observable physical quantities, aligning with Shepard's ecological representationalism and broader evolutionary perspectives (Gibson, 1979; Godfrey-Smith, 1993).

An essential feature of the Patchwork Approach is its ability to use finite experiences to predict novel events. For example, a child learning to bat in little league does not need to have seen every possible pitch to predict the time-to-contact of a ball (Figure 1a). Instead, the Patchwork Approach emphasizes interpolation—the ability to integrate across structured experiences within the range of prior encounters. This allows the child to make accurate judgments about the trajectory of a ball based on prior pitches with similar dynamics. As individuals accumulate a broader range of experiences, their representational space refines, enabling more accurate predictions and generalizations. This refinement arises not from the sheer quantity of experiences, but from exposure to a diverse array of physical systems and quantities. Consequently, sensorimotor predictions about physical interactions can emerge naturally without requiring firsthand experience of every situation, though the range and accuracy of these predictions are constrained by an individual's experiences.

Topological Structure of the Patchwork Approach

The Patchwork Approach to intuitive physics can be effectively conceptualized using topological principles that inform how experiences are added, weighted, and integrated into a coherent structure. This section delves into the topological framework that underpins this approach, providing a foundation for understanding how sensory experiences shape our physical representations.

By using a topological framework, the Patchwork Approach can effectively model how new experiences are integrated into our understanding of physical interactions while accounting for their relative influence based on their distance from established experiences. This relationship captures the intuitive idea that experiences far from the norm are less likely to be meaningfully integrated, thus enhancing our understanding of the complexities inherent in sensorimotor experiences.

The process of weighting experiences in the manifold reflects a dynamic transformation that accounts for the nature of experiences. As new stimuli are added to the psychological space, the manifold can stretch or shear, altering its geometry to ensure that predictions remain valid within a relevant range of experiences. This topological adjustment allows for accurate representations of physical relationships while integrating novel information.

By emphasizing the interplay between the manifold's shape and the weights assigned to experiences, we can better understand how perception navigates the complexities of physical interactions. This approach highlights that while distant or unfamiliar experiences may have a reduced influence, they still contribute to the evolving structure of the manifold, providing potential pathways for integrating new knowledge into our intuitive physics.

Adding Experiences In the context of the Patchwork Approach, experiences are added to the manifold based on observed physical interactions. Each point corresponds to a specific instance of experienced stimuli characterized by relevant physical features. The addition of an experience can be formalized as follows:

$$E_i = (f_1, f_2, \ldots, f_n),$$

where E_i represents the i-th experience, and $f_1, f_2, \ldots, f_n$ are the physical features associated with that experience. As new experiences are encountered, they are integrated into the manifold, which expands to accommodate this additional information.

Weighting Experiences In the Patchwork Approach, the weighting of experiences is a crucial mechanism governing how new experiences are integrated into the manifold of physical representations. This weighting is fundamentally based on the spatial relationships among experiences, reflecting the topological structure of the manifold and allowing it to adapt to new information. The weight assigned to each experience, E_i, is influenced by its distance from the nearest existing experience in psychological space (Figure 1b) and can be mathematically defined as:

$$E_{\text{nearest}} = \arg\min_{E_j \in \varepsilon} d(E_i, E_j),$$

where E_{nearest} represents the closest experience to E_i within ε, the set of all experiences in the manifold (excluding E_i itself). As the distance from a new experience increases, the manifold must stretch to incorporate it, leading to a reduced weighting for that experience(Figure 1c), :

$$S = d(E_i, E_{\text{nearest}}).$$

Here, S represents the shear in the manifold due to the addition of the new point. Consequently, the weight assigned to the new experience can be expressed as:

$$w_i = \frac{1}{(S+c)^{pd}},$$

where c is a small constant included to prevent division by zero and ensure stability in weight calculations. This relationship ensures that as the distance increases, the weight decreases, thereby diminishing the influence of experiences that are far removed from established clusters.

In this context, when the distance $d(E_i, E_{\text{nearest}})$ is small, the shear S is also small, resulting in a higher weight w_i. This indicates that the new experience is more relevant to the manifold's shape and fits in with previous experiences, ε. In contrast, when $d(E_i, E_{\text{nearest}})$ is significantly larger, shear increases, leading to a lower weight. Thus, the farther away the new experience, the less it influences the manifold, reflecting its reduced relevance (see Figure 2).

Manifold Surface Structure The surface of the manifold serves as a functional representation of these relationships, reflecting the underlying dynamics of physical interactions and allowing interpolation between experiences and predictions based on their weighted contributions. This process emphasizes the importance of the collective influence of nearby experiences in shaping the manifold structure, ultimately guiding our understanding and predictions of physical interactions.

A key feature of the Patchwork Approach is the local nature of its computations. Instead of requiring a global computation across the entire manifold, the approach relies on simple local computations convolved throughout the feature space (Rumelhart & McClelland, 1986). As the manifold surface is constructed incrementally, each new experience influences only its immediate neighborhood, with the manifold's slope changing non-uniformly. This local weighting mechanism ensures that predictions are sensitive to the clustering of experiences, with closer experiences having a stronger influence and more distant ones a diminishing effect. This approach is computationally efficient and biologically plausible, as it mirrors the localized distributed processing of neural systems (Friston, 2010).

The general shape of the manifold is formed by the arrangement of experiences as points in an n-dimensional space, where each dimension corresponds to a specific physical feature. This spatial configuration not only captures the individual characteristics of experiences, but also reflects the relationships between these features, allowing interpolation between these points and facilitating predictions based on their weighted contributions.

As new experiences are gained, the manifold undergoes geometric transformations, such as stretching or shearing—to incorporate this additional information while preserving the coherence of the underlying structure. These transformations enable the manifold to adapt dynamically, ensuring that it accurately represents the evolving relationships defined by the physical features of the perceptual experiences. The manifold M can be defined as a continuous surface that interpolates the experiences in the feature space:

$$M = \{(f_1, f_2, \ldots, f_n) \in \mathbb{R}^n : \forall E_i \in \varepsilon, M(f_1, f_2, \ldots, f_n)\},$$

where each f_i represents a specific dimension corresponding to a physical feature of the experiences (e.g., initial height, velocity, or mass).

The surface of the manifold is defined by a function $F(P)$

that combines the contributions of each experience based on their weights:

$$F(P) = \sum_{i=1}^{N} w_i f(E_i),$$

where $F(P)$ represents the value of the manifold's surface at point P, N is the total number of experiences, w_i is the weight assigned to each experience E_i, and $f(E_i)$ is a function that defines the contribution of each experience to the surface, based on its physical features. As new inputs are encountered, the manifold's surface dynamically adjusts based on the weighted relationships among experiences.

The flexibility inherent in the manifold's structure enables it to adapt to varying levels of dimensionality. It can stretch and reshape in response to new experiences, thereby reflecting the complexities of sensorimotor experiences and reinforcing the notion that sensorimotor prediction is a product of both historical interactions and current perceptual inputs. This process is akin to the consolidation of sensory experiences, similar to how perceptual learning consolidation stabilizes sensory information over time (Sasaki et al., 2010).

Gradient Computation In the Patchwork Approach, the gradient of the manifold is essential for understanding the relationships between the various dimensions that characterize sensory experiences. This gradient indicates how changes in one physical feature relate to changes in others, capturing the interconnectedness of the dimensions within a single experience. By analyzing the gradient, we can identify how different features influence the overall representation of physical interactions, allowing for nuanced predictions and understanding of behavior based on past experiences. The gradient of the manifold at a specific point P can be mathematically defined as follows:

$$\nabla F(P) = \left(\frac{\partial F}{\partial f_1}, \frac{\partial F}{\partial f_2}, \ldots, \frac{\partial F}{\partial f_n} \right),$$

where $F(P)$ represents the value of the manifold's surface at point P. Each partial derivative, such as $\frac{\partial F}{\partial f_1}$, quantifies the sensitivity of the manifold's value to changes in the corresponding feature, indicating the degree to which variations in that dimension impact the overall representation. By assessing these partial derivatives, we gain insight into which features are most influential in shaping the manifold and how tightly coupled these features are within the physical interactions being represented.

By analyzing these gradients, we can identify how experiences, which serve as discrete instances of experience, inform our understanding of physical laws. Specifically, the gradient reveals the direction and magnitude of influence among the dimensions, facilitating the extrapolation of general principles from specific experiences. This capability allows us to formulate predictions about novel physical scenarios based on the established relationships encapsulated within the manifold. In essence, the gradient acts as a bridge between the discrete experiences represented by the experiences and the broader laws of physics, enhancing our ability to anticipate how physical interactions will unfold in various contexts.

To illustrate the predictive power of the gradient, consider the specific case of catching a ball, where the relationship between velocity in depth, height of the throw, and the projectile's location in depth is critical (Figure 1a). As the velocity of a thrown ball increases, there is a pronounced increase in the final location of the object, reflecting a strong covariation between these two dimensions. In contrast, variations in height normally exert a comparatively smaller influence on location in depth. This differential sensitivity can be mathematically captured by the gradient of the manifold, which reveals the extent to which changes in velocity in depth affect the predicted final location in depth.

For example, when evaluating the gradient at a specific point P, the partial derivatives $\frac{\partial F}{\partial v_z}$ and $\frac{\partial F}{\partial h}$ quantify how sensitive z is to changes in velocity and height, respectively. Additionally, we can define a partial derivative for location in depth itself, $\frac{\partial z}{\partial v_z}$ and $\frac{\partial z}{\partial h}$, which provides insight into how these changes impact the projectile's location. The collection of these partial derivatives represents the *internalized laws of physics*, effectively summarizing how sensory information relates to physical predictions.

While the gradient facilitates interpolation—making predictions within the established range of experiences—extrapolation beyond that range becomes uncertain. The reliability of predictions outside the experience range diminishes, as they rely on the established relationships defined by nearby points. Therefore, the gradient serves as a crucial tool in modeling intuitive physics by providing a framework for understanding how sensory experiences inform our predictions of physical interactions within the boundaries of observed data.

Evaluation of the Patchwork Approach

This section evaluates the Patchwork Approach through two empirical tests, demonstrating its ability to model human perceptual predictions.

Empirical Test 1: Gravity Perception

This test applies the Patchwork Approach to data from Deeb & Domini (2024), examining the perception of a projectile's final location under varying gravity conditions. Participants viewed a projectile motion event and adjusted a probe to match the perceived depth position in a virtual reality environment. The Patchwork Approach is used here to demonstrate how sensory cues, such as depth and velocity, are integrated to implicitly internalize gravity, allowing for accurate predictions without direct modeling of gravitational forces. Results suggest that the Patchwork Approach provides a more adaptable framework for understanding human perception of gravity and its relationship to depth perception.

Data Generation and Manifold Construction To model the perceptual judgments in Deeb & Domini's (2024) Experiment 1, a manifold was constructed with 5000 randomly generated stimuli to capture the relationship between three behaviorally relevant features: height h, vertical velocity v_z, and depth position z (fig. 2a). The model used vertical velocities ranging from 0.1 to 1.8 m/s and heights ranging from 0.025 to 1m. To simulate the influence of sensory experiences on the manifold, proximity-based weighting was applied to each data point based on its spatial distance to others in the feature space.

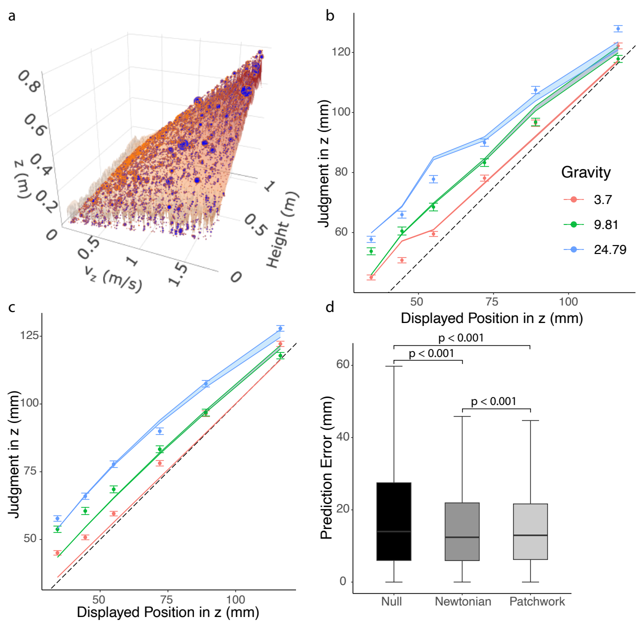

Figure 2: a) The manifold of projectile motion from Empirical Test 1, with experiences coded in blue, shown across the axes of distance (z), velocity in depth v_z, and height (h). (b) The Patchwork model fit to the data, with model predictions represented by the curved lines and the empirical data from Experiment 1 of Deeb & Domini (2024), including ±1 SEM error bands. The vertical axis shows the perceived depth location, with 0 mm representing the launch position, and the horizontal axis displays the displayed position. Color coding reflects different gravity conditions. (c) Newtonian ground truth fit with biased velocity in depth. (d) Boxplot comparison of prediction errors for the Null, Newtonian, and Patchwork models, showing significant differences (paired t-test, two-tailed).

Cue Combination The constructed manifold was used to predict the final z position of the projectile under varying gravity conditions. To match the experimental setup from Deeb & Domini (2024), a slice of the manifold was extracted corresponding to the specific height $h = 67.83$ cm used in the experiment. Within this slice, the previously calculated positions in depth were retrieved for various vertical velocities, which were aligned with the specific range of velocities observed in the study. These predictions were then used as part of the cue combination process with the sensory information (fig. 2b).

The Patchwork Approach itself is agnostic to the specific method of cue combination, as its focus is on explaining how sensorimotor experiences are consolidated into representations within topological space. The method employed in this example uses a weighted average model, which was selected for its simplicity and consistency with the approach used in the original study by Deeb & Domini (2024) (fig. 2c). This model integrates the manifold predictions z_p and sensory data z_s using a weighted average model:

$$\hat{z} = z_p w_p + z_s w_s,$$

where w_p and w_s represent the weights for the manifold prediction z_p and the sensory input z_s, respectively. This method allows for an effective integration of the manifold predictions and sensory inputs, with the weights dynamically adjusting based on the relationship between the data points.

Statistical Analysis A paired-samples t-test showed a significant difference in prediction errors between the Patchwork model ($M = 20.67$, $SD = 12.45$) and the previous model ($M = 21.95$, $SD = 12.85$); $t(7919) = -6.66$, $p < 0.001$. The mean error difference was -0.43 mm (95% CI $[-0.56, -0.31]$) with the Patchwork model consistently outperforming the previous model and the null model wherein gravity is not a cue to depth (fig. 2d). Additionally, correlations between model predictions and human responses revealed that the Patchwork model ($r = 0.80$) aligned more closely with human data compared to the previous model ($r = 0.75$). These findings indicate that the Patchwork model provides more accurate and human-like predictions of depth position.

Empirical Test 2: Interpolation in Angular Kinematics

This test evaluates the Patchwork Approach's ability to generalize to novel conditions by interpolating untrained aiming angles in a deflection angle prediction task. Using data from Deeb et al. (2021), which examined how Newtonian predictions correct noisy perception in collision events, the Patchwork model was tested on its ability to predict deflection angles for aiming angles not included in its training data. The model successfully interpolated to untrained aiming angles, demonstrating its capacity to generalize based on learned relationships between aiming and deflection angles. While the Patchwork model was not explicitly trained on the tested angles, it performed comparably to the Newtonian ground truth model.

Data Generation and Manifold Construction The experimental setup from Deeb et al. (2021) was used, where aiming angles of ±10° resulted in deflection angles of approximately ±23°. To test generalization, the training data included aiming angles within a wider range of ±15° but excluded a large range of aiming angles between ±8° and ±12°, specifically omitting the ±10° used in the original study. Distances between the cue and target objects varied between 80mm and 120mm, encompassing the 100mm distance used in Deeb et al. (2021). The model was trained on 5000 randomly generated, physically consistent collision events. Proximity-based weighting was applied to both aiming angles and distances in the feature space, as described in the first test.

Manifold Slicing and Prediction A slice of the manifold was extracted at 100mm, matching the distance used in Deeb et al. (2021). The model interpolated deflection angles for the untrained aiming angles ($\pm 10^\circ$), accurately predicting a deflection angle of approximately $\pm 21^\circ$.

Cue Combination and Optimization Predictions for the $\pm 10^\circ$ aiming angle at 100mm were integrated into the Maximum Likelihood Estimate (MLE) model from Deeb et al. (2021). The deflection angle was calculated as a weighted combination of sensory estimates and Patchwork predictions, with the weights based on the variance of sensory noise and the Newtonian prediction.

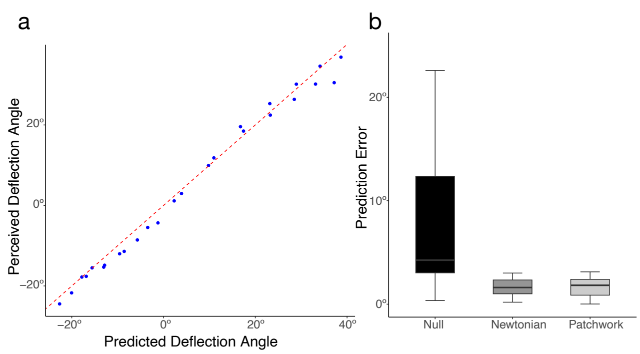

Figure 3: a) Prediction of deflection angles using the Patchwork model compared to human data from the angular kinematics study by Deeb et al. (2021). The red dashed line represents the unity line where model predictions match the actual deflection angles. (b) Box plot comparing the prediction errors between the Null, Newtonian, and Patchwork models.

Statistical Analysis The Patchwork model showed a slight improvement in RMSE compared to the Newtonian model (RMSE = 1.93 vs. 2.10), highlighting its ability to generalize to untested conditions (fig. 3a). There was no significant difference in prediction error between the Newtonian and Patchwork models ($t(25) = -1.02$, $p = 0.317$), though both significantly outperformed the null model, which relied solely on sensory information (Newtonian vs. Null: $t(25) = -4.22$, $p < 0.001$; Patchwork vs. Null: $t(25) = -4.26$, $p < 0.001$) (fig. 3b). These results demonstrate the Patchwork model's ability to interpolate and generalize to novel conditions by leveraging physical relationships.

Discussion

The Patchwork Approach was tested on two different experimental datasets, each addressing distinct aspects of physical perceptual prediction. In the first experiment, which involved projectile motion, the goal was to determine whether the Patchwork model could improve upon the previously established Newtonian-based model. The results demonstrate that the Patchwork model provides a better fit, reflecting its ability to accurately predict the final depth positions of the projectile. In the second experiment, the focus shifted to testing the model's ability to generalize to novel aiming angles, specifically those that were not included in the training set. The Patchwork model was able to interpolate the deflection angles for these untested conditions, performing as well as the Newtonian ground truth model.

The Patchwork Approach excels in modeling intuitive physics by interpolating based on encountered experiences, as demonstrated by its superior performance in predicting projectile motion and interpolating deflection angles for untrained conditions. The model internalizes physical relationships from sensorimotor experiences, applying them to novel scenarios with efficiency. It relies on local computations, where each experience influences only its immediate neighborhood, mimicking the localized, distributed processing of neural systems. However, the model's focus on sensorimotor representations limits its applicability to more abstract reasoning tasks, which may lie outside the realm of sensorimotor processing. This limitation could be addressed by integrating frameworks like the Intuitive Physics Engine (IPE). However, unlike traditional frame-prediction models, which attempt to predict the next frame in a physical scene, the Patchwork Approach focuses on the behaviorally relevant features of dynamic events. This focus is grounded in empirical data related to human perception, specifically sensorimotor interactions with the physical world. For example, recent studies have shown that adults, although poor at predicting the trajectory of objects after release from a pendulum, can accurately predict the landing location of these objects (Smith & Vul, 2013). This suggests that humans rely on behaviorally relevant cues, rather than trying to predict each frame of motion.

Although the Patchwork Approach's current focus is on perceptual predictions in the domain of intuitive physics, there is potential for extending it to other dynamic percepts. One promising direction is the study of social perception, where there are regular visual precursors to higher-order latent variables, such as understanding others' intentions or actions (Isik et al., 2018, 2020). The Patchwork Approach's capacity to capture feature-based correlations may provide a strong framework for modeling social dynamics. Additionally, incorporating computer vision models could enhance the Patchwork Approach's predictions by enabling direct processing of visual inputs, further expanding its utility in real-world perception.

In summary, the Patchwork Approach establishes a flexible, experience-driven framework for physical predictions, capable of generalizing to novel scenarios and interpolating dynamics from limited observations without exhaustive training. Nonetheless, future work should extend it to more complex physical systems and integrate multimodal transformations—such as how haptic and visual regularities jointly guide perception and action (White, 2012).

References

Badler, J., Lefèvre, P., & Missal, M. (2010). Causality attribution biases oculomotor responses. *The Journal of Neuroscience*, *30*(31), 10517–10525. doi: 10.1523/JNEUROSCI.1733-10.2010

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. doi: 10.1073/pnas.1306572110

Bremner, A. J., Holmes, N. P., & Spence, C. (2012). The development of multisensory representations of the body and of the space around the body. In A. J. Bremner, D. J. Lewkowicz, & C. Spence (Eds.), *Multisensory development* (pp. 113–136). Oxford University Press. doi: 10.1093/acprof:oso/9780199586059.003.0005

Deeb, A.-R., Cesanek, E., & Domini, F. (2021). Newtonian predictions are integrated with sensory information in 3d motion perception. *Psychological Science*, *32*(2), 280–291. doi: 10.1177/0956797620966785

Deeb, A. R., & Domini, F. (2024). Embeddedness of earth's gravity in visual perception. *Journal of Vision*, *24*(11), 4. doi: 10.1167/jov.24.11.4

Deeb, A. R., & Domini, F. (2025). [the online processing of dynamics]. *Cognition*. (Under review)

Deeb, A. R., & Fischer, J. (2025). [velocity—not perceived as such: The role of perceived mass on velocity estimation]. *[In preparation]*.

Deeb, A.-R., Silva, A., & Liu, Z. (2022). Causality modulates perception of apparent motion stimuli. *Vision Research*, *201*(10), 108120. doi: 10.1016/j.visres.2022.108120

Fischer, J., Mikhael, J. G., Tenenbaum, J. B., & Kanwisher, N. (2016). Functional neuroanatomy of intuitive physical inference. *Proceedings of the National Academy of Sciences*, *113*(34), E5072–E5081. doi: 10.1073/pnas.1610344113

Friston, K. (2010). The free-energy principle: a unified brain theory? *Nature Reviews Neuroscience*, *11*, 127–138. doi: 10.1038/nrn2787

Gibson, J. J. (1979). *The ecological approach to visual perception*. Houghton, Mifflin and Company.

Godfrey-Smith, P. (1993). Functions: Consensus without unity. *Pacific Philosophical Quarterly*, *74*, 196–208. doi: 10.1111/j.1468-0114.1993.tb00358.x

Isik, L., Mynick, A., Pantazis, D., & Kanwisher, N. (2020). The speed of human social interaction perception. *NeuroImage*, *215*, 116844. doi: 10.1016/j.neuroimage.2020.116844

Isik, L., Tacchetti, A., & Poggio, T. (2018). A fast, invariant representation for human action in the visual system. *Journal of Neurophysiology*, *119*(2), 631–640. doi: 10.1152/jn.00501.2017

Kaiser, M. K., Proffitt, D. R., & McCloskey, M. (1985). The development of beliefs about falling objects. *Perception & Psychophysics*, *38*(6), 533–539. doi: 10.3758/BF03207062

Kaiser, M. K., Proffitt, D. R., Whelan, S. M., & Hecht, H. (1992). Influence of animation on dynamical judgments. *Journal of Experimental Psychology: Human Perception and Performance*, *18*(3), 669–689. doi: 10.1037/0096-1523.18.3.669

Kim, S.-H., Feldman, J., & Singh, M. (2013). Perceived causality can alter the perceived trajectory of apparent motion. *Psychological Science*. doi: 10.1177/0956797612458529

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), 952–966. doi: 10.1037/0278-7393.19.4.952

Mason, R. A., & Just, M. A. (2016). Neural representations of physics concepts. *Psychological Science*, *27*(6), 904–913. doi: 10.1177/0956797616641941

McCloskey, M., Caramazza, A., & Green, B. (1980). Curvilinear motion in the absence of external forces: naive beliefs about the motion of objects. *Science (New York, N.Y.)*, *210*(4474), 1139–1141. doi: 10.1126/science.210.4474.1139

McIntyre, J., Zago, M., Berthoz, A., & Lacquaniti, F. (2001). Does the brain model newton's laws? *Nature Neuroscience*, *4*(7), 693–694. doi: 10.1038/89477

Michotte, A. (1963). *The perception of causality*. Basic Books.

Mitko, A., & Fischer, J. (2023). Do striking biases in mass inference reflect a flawed mental model of physics? *Journal of Experimental Psychology: General*, *152*(9), 2636–2650. doi: 10.1037/xge0001399

Mitko, A., Navarro-Cebrián, A., Cormiea, S., & Fischer, J. (2024). A dedicated mental resource for intuitive physics. *iScience*, *27*(1), 108607. doi: 10.1016/j.isci.2023.108607

Monache, S. D., Lacquaniti, F., & Bosco, G. (2019). Ocular tracking of occluded ballistic trajectories: Effects of visual context and of target law of motion. *Journal of Vision*, *19*(4), 13. doi: 10.1167/19.4.13

Montagne, G., Laurent, M., Durey, A., & Bootsma, R. (1999). Movement reversals in ball catching. *Experimental Brain Research*, *129*(1), 87–92. doi: 10.1007/s002210050939

Peters, M. A., Balzer, J., & Shams, L. (2015). Smaller = denser, and the brain knows it: Natural statistics of object density shape weight expectations. *PLoS ONE*, *10*(3), e0119794. doi: 10.1371/journal.pone.0119794

Piloto, L. S., Weinstein, A., Battaglia, P., et al. (2022). Intuitive physics learning in a deep-learning model inspired by developmental psychology. *Nature Human Behaviour*, *6*, 1257–1267. doi: 10.1038/s41562-022-01394-8

Rolfs, M., Dambacher, M., & Cavanagh, P. (2013). Visual adaptation of the perception of causality. *Current Biology*, *23*(3), 250–254. doi: 10.1016/j.cub.2012.12.017

Ross, H. E. (1969). When is a weight not illusory? *The Quarterly Journal of Experimental Psychology*, *21*(4), 346–355. doi: 10.1080/14640746908400230

Rumelhart, D. E., & McClelland, J. L. (1986). *Parallel distributed processing: Explorations in the microstructure of cognition. volume 1. foundations*. MIT Press.

Runeson, S., & Frykholm, G. (1983). Kinematic specification of dynamics as an informational basis for person-and-action perception: Expectation, gender recognition, and deceptive intention. *Journal of Experimental Psychology: General*, *112*(4), 585–615. doi: 10.1037/0096-3445.112.4.585

Sanborn, A. N., Mansinghka, V. K., & Griffiths, T. L. (2013). Reconciling intuitive physics and newtonian mechanics for colliding objects. *Psychological Review*, *120*(2), 411–437. doi: 10.1037/a0031912

Sasaki, Y., Nanez, J. E., & Watanabe, T. (2010). Advances in visual perceptual learning and plasticity. *Nature Reviews Neuroscience*, *11*(1), 53–60. doi: 10.1038/nrn2737

Shepard, R. N. (1987). Toward a universal law of generalization for psychological science. *Science*, *237*(4820), 1317–1323. doi: 10.1126/science.3629243

Shepard, R. N. (1994). Perceptual-cognitive universals as reflections of the world. *Psychonomic Bulletin & Review*, *1*(1), 2–28. doi: 10.3758/BF03200759

Siegler, I. A., Bardy, B. G., & Warren, W. H. (2010). Passive vs. active control of rhythmic ball bouncing: The role of visual information. *Journal of Experimental Psychology: Human Perception and Performance*, *36*(3), 729–750. doi: 10.1037/a0016462

Smith, K., & Vul, E. (2013). Consistent physics underlying ballistic motion prediction. In *Proceedings of the 35th annual meeting of the cognitive science society.*

Todd, J. T., & Warren, W. H. (1982). Visual perception of relative mass in dynamic events. *Perception*, *11*(3), 325–335. doi: 10.1068/p110325

Walker, P., Francis, B. J., & Walker, L. (2010). The brightness-weight illusion. *Experimental Psychology*, *57*(6), 462–469. doi: 10.1027/1618-3169/a000057

Wang, H., Jedoui, K., Venkatesh, R., et al. (2024). Probabilistic simulation supports generalizable intuitive physics. In *Proceedings of the 46th annual meeting of the cognitive science society.*

White, P. A. (2012). The experience of force: The role of haptic experience of forces in visual perception of object motion and interactions, mental simulation, and motion-related judgments. *Psychological Bulletin*, *138*(4), 589–615. doi: 10.1037/a0025587