## Title
Physical and Causal Judgments for Object CollisionsDepend on Relative Motion



## Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 40(0)

## Authors
Kubricht, James R
Lu, Hongjing

## Publication Date
2018

# Physical and Causal Judgments for Object Collisions Depend on Relative Motion

**James R. Kubricht**[1] **Hongjing Lu**[1,2]
Department of Psychology[1] Department of Statistics[2]
University of California, Los Angeles

**Abstract**

Human judgments about the physical attributes of—and causal relationship between—two colliding objects have been studied extensively over the past seventy years. Recent computational evidence suggests that judgments about the mass ratio of two colliding objects, as well as their perceived causal relation, can be explained by a coherent framework based on a Newtonian physical model and probabilistic inference resulting from noisy observations of object movements. However, it remains unclear how the physical and causal reasoning systems interact with the motion perception system when forming these judgments. The current study aims to examine whether high-level judgments are guided by object motion represented as relative motion with reference to a moving background, or as absolute motion with reference to a stationary position in the world. Both experimental evidence and model simulation results support the notion that physical and causal inference in object collisions depend on relative motion rather than absolute motion.

**Keywords:** Intuitive physics; causality; mass judgment; reference frame; Bayesian inference

## Introduction

Over the past seventy years, researchers have examined how human inferences about the attributes of—and causal relationship between—colliding objects vary according to spatiotemporal properties in observed displays (Cohen, 2006; Gilden & Proffitt, 1994; Leslie, 1982; Michotte, 1963; Natsoulas, 1961; Runeson, 1983; Runeson, Juslin, & Olsson, 2000; Saxe & Carey, 2006; Schlottmann & Anderson, 1993; Scholl & Nakayama, 2002; Todd & Warren Jr., 1982; White, 2006). In a typical launching event, an initially moving disc (motor object; or Object $A$) collides with an initially stationary one (projectile object; or Object $B$) and causes it to move forwards in the direction that it is pushed. Physically, the motor object interacts with the projectile object by imparting its momentum upon it; the sum of the motor and projectile objects' momentum remains constant over time: i.e., the principle of conservation of momentum.

Michotte (1963) found that when people observe launching events, they report an immediate and irresistible impression that the motor object *causes* the projectile to move forwards. However, participants' causal ratings were consistently attenuated when either (1) a spatial gap was placed between the colliding objects; (2) the projectile object's movement was delayed; or (3) the projectile object moved faster/slower than the motor object following impact (see Figure 1). These findings indicate that causal impressions are highly sensitive to the spatiotemporal characteristics of observed events. However, Runeson (1983) later pointed out that causal impressions were too subjective to reliably measure and instead turned his attention towards relative mass judgments. Following Gibson's (1966) doctrine of direct perception, he theorized that if people reason according to the principle of conservation of momentum, their judgments about which of two colliding objects is heavier should solely depend on their unbiased and accurate estimates of each object's pre- and post-collision velocity.

Runeson's predictions were subsequently tested by Todd and Warren (1982) who found that people are instead consistently biased towards reporting that the motor object is heavier (i.e., the motor object bias). Moreover, people are more susceptible to this bias when the objects are relatively inelastic: e.g., deformable vs. rigid balls. To explain these findings, researchers posited that observers form judgments according to simplified heuristic rules based on salient perceptual cues: e.g., which object moves faster after impact and the degree to which each object deflects off of the other (Gilden & Proffitt, 1994; Runeson et al., 2000). Although the heuristic approach qualitatively explains trends in the reported behavioral data, a more recent approach has demonstrated that people *do* appear to reason about relative mass in accordance with the principle of conservation of momentum, given that their perception is prone to error and prior beliefs about informative physical variables are held: i.e., the *noisy Newton* hypothesis (Sanborn, Mansinghka, & Griffiths, 2013; Sanborn, 2014).

### The Noisy Newton Framework

The noisy Newton framework was proposed to explain people's predictions about dynamic physical situations without the implementation of arbitrary heuristic rules. It has been employed across a wide range of physical domains, ranging from the movement of non-solid substances to causal reasoning through counterfactual simulation (see Kubricht, Holyoak, & Lu, 2017 for a review). The framework supposes that people possess an intuitive physics "engine" (Battaglia, Hamrick, & Tenenbaum, 2013) encoded in neural circuitry (Fischer, Mikhael, Tenenbaum, & Kanwisher, 2016) which approximately emulates the laws of physics to simulate spatially represented variables forwards in time (see Battaglia, Pascanu, Lai, & Rezende, 2016; Chang, Ullman, Torralba, & Tenenbaum, 2016; Grzeszczuk, Terzopoulos, & Hinton, 1998 for a computational approach). Moreover, since people's observations are inherently noisy, inferred estimates of observable variables are consistently biased towards prior expectations.

In the case of object collisions, Sanborn et al.'s (2013) implementation of the noisy Newton framework adopts the generic prior in motion perception (i.e., favoring slow motion in object motion) and a likelihood function which compares

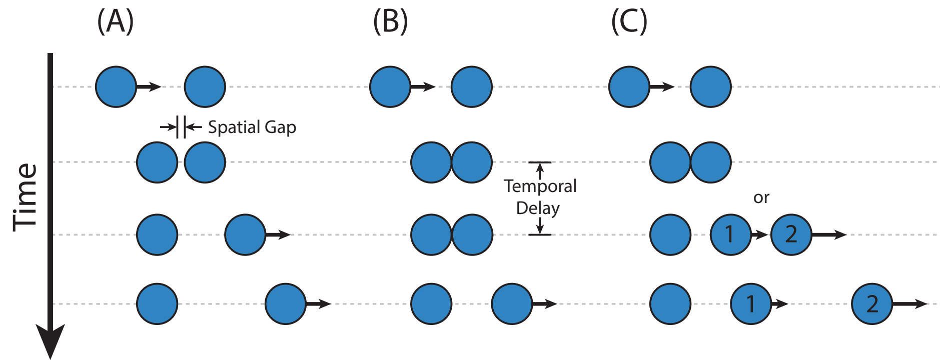


Figure 1: Three common manipulations to spatiotemporal properties of a collision event display in causal perception tasks. The dashed grey lines correspond with different time points during the collision event, and the arrows attached to the discs indicate their magnitude and direction of movement (i.e., velocity). (A) When a spatial gap is placed between the inside edges of two colliding discs at the moment of impact, perceived causality diminishes. (B) Introducing a pause (temporal delay) between when the motor object stops and the projectile object begins moving also attenuates perceived causality. (C) The numbers (1 or 2) indicate two possible outcomes; the projectile object moves either slower than the motor object following collision (Outcome 1) or it moves faster (Outcome 2). People report greater causal impressions after observing Outcome 1 relative to Outcome 2.

the observed velocity with the derived velocity from a physical model. The noisy Newton approach explains the motor object bias and predicts larger biases for relatively inelastic collisions. In addition, the noisy Newton model can be extended to account for causal judgments by comparing how well a noisy Newtonian model explains observations compared with a non-physical model (Sanborn et al., 2013; see Appendix for model details).

Importantly, the noisy Newton framework does not specify how observable input variables should be represented: e.g., in a dynamic stimulus, object motion can be represented as relative motion with reference to a moving background, or as absolute motion with reference to a stationary position in the world. Furthermore, the physical inference may vary depending on whether the reasoning system adopts relative or absolute motion signals. For instance, imagine that you are looking out of the left window of a resting train and you see a vehicle move from your left periphery towards a second vehicle parked on the road nearby. The two vehicles collide, and the second vehicle correspondingly moves in the direction that it was pushed. You might get an impression that the two vehicles were equally heavy; but what if the train was traveling in the same direction as the initially moving vehicle? What if it was traveling in the opposite direction? Would your judgment about the weight of the two vehicles change? Would you be equally likely to report that the first vehicle had *launched* the second one forwards? Motion perception studies have shown that humans can perceive both relative and absolute motion with different degrees of sensitivity (Smeets & Brenner, 1994). For cognitive tasks probing the ability of physical and causal reasoning, it is important to understand what perceptual variables are selected and used for high-level judgments.

The central effort of the current experiments is to examine what motion information is extracted from visual inputs for physical and causal judgments. Previous work on object collision judgments have exclusively used stationary backgrounds, providing no distinction between relative and absolute motion. However,in daily life, perceived landmarks are constantly moving across our visual field as we move through—and interact with—the environment. In such cases, representing motion relative to those moving landmarks could provide a different explanation of physical dynamics than absolute motion does. Across two experiments, we (1) measured human performance in physical and causal judgment tasks when viewing object collisions on a moving background, and (2) compared noisy Newton model predictions given absolute and relative motion inputs to test the hypothesis that humans encode relative motion when forming mass judgments and inferring causality.

## Experiment 1: Mass Judgments

The goal of the first experiment was to determine (1) whether a vertical background grid moving with or against the motion of two colliding objects influences mass ratio judgments; and (2) if so, whether the noisy Newton model for mass collisions with relative motion inputs can explain participants' performance.

### Participants

A total of 20 undergraduate students (14 female; Mean age = 21.2) were recruited from the University of California, Los Angeles (UCLA) Department of Psychology subject pool and were compensated with course credit.

### Materials and Procedure

Collision event videos were presented on a 19" Dell E198WFP LCD monitor with a refresh rate of 40 Hz at 1440 $\times$ 900 resolution. Videos were viewed at a distance of approximately 70 cm. In each video, an initially moving object (termed as motor object) collided with a stationary ($u_B = 0$) object (termed as projectile object). The pre-collision velocity of the motor object varied across eight values: $u_A = 1.9$, 2.3, 2.6, 3.0, 3.4, 3.7, 4.1, and 4.5 cm/sec. The final velocities of the two objects were determined by Newtonian principles

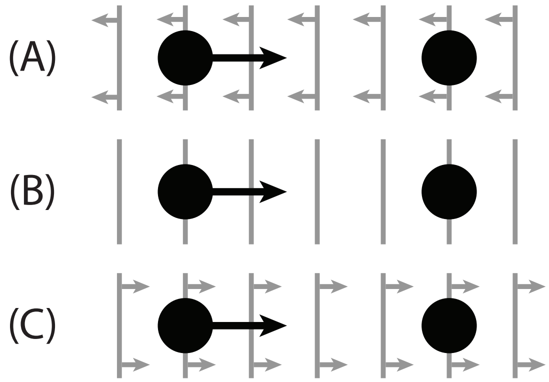


Figure 2: The three panels depict a collision event prior to impact where the motor object (left) travels towards the projectile object (right). The vertical gray lines indicate a background grid which either (A) moves leftward against the direction of the collision event, (B) remains stationary, or (C) moves rightward in the direction of the collision event. The black and gray arrows indicate the motor object and background grid velocities, respectively.

using a fixed restitution value of $e = 0.9$ and eight mass ratio values: $m_A/m_B$ = 1/3, 1/2, 2/3, 4/5, 5/4, 3/2, 2/1, 3/1. In each video, the background also either moved against the direction of the collision (leftward; -2 cm/sec), with the direction of the collision (rightward; 2 cm/sec), or it remained at rest (0 cm/sec; see Figure 2). These manipulations yielded 8 (motor speed) $\times$ 8 (mass ratio) $\times$ 3 (background movement) = 192 collision stimuli presented in a within-subjects design. Trials were presented in a randomized order, and no feedback was provided. Each stimulus video lasted 4 sec with impact occurring 2 sec into each collision event; the impact location was always at the center of the display. The collision videos were rendered using MATLAB Psychophysics Toolbox 3. The motor and projectile objects were depicted as black (RGB = 0 0 0) discs with 2.7 cm diameter. The vertical grid lines spanned the height of the screen (25.4 cm) and were colored gray (RGB = 150 150 150). Each line was 0.08 cm wide with a horizontal line separation of 2.7 cm.

Prior to the testing trials, participants were informed that they would be watching a series of videos where two discs interact with one another. They were told that there would be vertical lines behind the two discs in each display and that they would either move leftward/rightward or remain at rest. In each trial, participants viewed a collision video and then reported which of the two objects (left or right) they thought was heavier. Participants were provided with the opportunity to take two breaks which occurred 1/3 and 2/3 of the way through the experiment, which lasted approximately 20 minutes.

## Human Results

The proportion of participants choosing the motor object as appearing heavier in each mass ratio and background movement condition is displayed in the left panel of Figure 3. Participants' responses—either 0 or 1—were averaged across the pre-collision motor velocity ($u_A$) conditions prior to analysis. A two-way repeated measures ANOVA was conducted on the response proportions to determine whether mass ratio and background movement influenced mass judgments. Results from the analysis indicated a significant interaction between mass ratio and background movement, $F(14,6) = 7.37$, $p = .01$, indicating that the impact of background movement on mass judgments varied according to mass ratio. As evident in Figure 3, participants were more likely to report that the motor object was heavier than the projectile object when the background moved leftward against the direction of the collision (Figure 2A), and less likely when it moved rightward in the same direction (Figure 2C). In other words, the point of subjective equality (PSE; i.e., the mass ratio where each judgment is equally likely) occurred at a minimum mass ratio with leftward background movement, PSE = 0.62, a moderate ratio when the background was at rest, PSE = 0.70, and a maximum ratio with rightward background movement, PSE = 0.99. In the following section, we explore whether the noisy Newton model for mass ratio judgments can explain this behavioral trend.

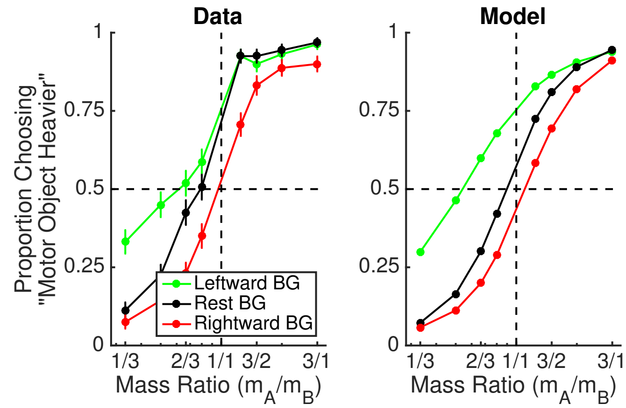


Figure 3: The measured (left panel) and model predicted (right panel) proportions of participants choosing the motor object as appearing heavier. Separate lines indicate whether the background (BG) moved leftward/rightward or remained at rest. Proportions are averaged across pre-collision velocity ($u_A$) conditions.

## Model Results

The noisy Newton model for mass ratio judgments (Sanborn et al., 2013) takes as input the velocities of the motor and projectile objects and outputs the likelihood of an observer choosing the motor object as appearing heavier. In the original noisy Newton model, the input of observed velocity is specified relative to a fixed point on the display; we will refer to these velocities as *absolute* velocities. This model can account for the motor object bias and predicts that the probability of choosing "motor object heavier" changes as a sigmoid function of the true mass ratio. The noisy Newton model with absolute velocity inputs is represented by the black curve in Figure 3 (right panel) and reveals a fit of $r^2(22) = 0.91$ (95% CI = [0.80, 0.95]). However, critically, the model's performance is not influenced by the presence and direction of background movement—nor do the model predictions differ—since the absolute velocity inputs do not change across the three background movement conditions. Therefore, the model predicts the same PSE in each background condition, PSE = 0.95.

Alternatively, the noisy Newton model can take as input the motor and projectile velocities specified *relative* to a moving point fixed to the background grid. The result is relatively large velocities when the background moves leftward and relatively small velocities when it moves rightward. Since observation noise in the noisy Newton model increases for larger velocity magnitudes, the influence of the slow motion prior is greatest in the leftward background condition and smallest in the rightward background condition. As shown by the separate curves in the right panel of Figure 3, the noisy Newton model with relative velocity inputs explains people's increasing bias towards reporting "motor object heavier" in the leftward versus rightward background movement condition. The model also provides a superior fit to human judgments, $r^2(22) = 0.97$ (95% CI = [0.91, 0.98]), and predicts human PSEs: Leftward PSE = 0.54, Rest PSE = 0.95, Rightward PSE = 1.09. The model results for Experiment 1 used the same parameters reported in Sanborn et al. (2013): i.e., $\sigma = 2$, $k_v = .1$, $w_v = .15$.

## Experiment 2: Causal Ratings

Our first experiment showed that the magnitude of the motor object bias depends on the background movement direction in a collision event. The noisy Newton model with relative motion inputs accounts for human mass ratio judgments well across a range of testing conditions. The purpose of the second experiment was to determine whether the same background manipulation affects perceived causality, and whether the noisy Newton model can account for human performance.

### Participants

A total of 29 undergraduate students (20 female; Mean age = 20.5) were recruited from the University of California, Los Angeles (UCLA) Department of Psychology subject pool and were compensated with course credit.

### Materials and Procedure

The apparatus was the same as in Experiment 1. The stimuli in Experiment 2 were also the same as previously indicated, except two differences: (1) the motor object was always stationary after impact (i.e., $v_A = 0$ cm/sec) and (2) the motor object moved comparatively faster: $u_A$ = 6, 11, and 15 cm/sec. Instead of using mass ratio, restitution, and each object's pre-collision velocity to determine their post-collision velocities in each trial, the ratio of the motor object's pre-collision velocity to the projectile object's post-collision velocity (see Figure 1C) was directly manipulated across trials: $u_A/v_B$ = 0.5, 0.7, 1, 1.4, 2. In addition, a temporal delay (see Figure 1B) was placed between the moment of impact and the projectile object's initial movement: $t$ = 0, 70, 140, 210, 280 msec. These manipulations yielded 3 (motor speed) × 5 (velocity ratio) × 5 (temporal delay) × 3 (background movement) = 225 collision stimuli presented in a within-subjects design. The trials were presented in a randomized order and no feedback was provided. The experiment lasted approximately 30 minutes.

Participants began the experiment by viewing a set of instructions informing them that they would be viewing videos of two (equally heavy) discs in motion. Once again, they were informed that there would be vertical grid lines behind the discs that would move leftward/rightward or remain at rest. Following each video, participants were asked, "Did the left object *launch* the right object?" and responded on a scale from 1 (Definitely No) to 9 (Definitely Yes) with a middle rating of 5 (Unsure).

### Human Results

As in the previous experiment, we averaged individual participants' ratings across the three trials with different pre-collision motor velocity ($u_A$). Mean ratings in each of the temporal delay, background movement, and velocity ratio conditions are displayed in the top panels of Figure 4. A three-way repeated measures ANOVA was conducted on the mean causal ratings with three within-subjects factors. There was a significant two-way interaction between temporal delay and velocity ratio, $F(16,13) = 2.90$, $p = .03$, indicating that the impact of velocity ratio on causal ratings depended on the magnitude of temporal delay. The three-way interaction and remaining two-way interactions were not statistically significant.

The impact of velocity ratio on causal ratings was examined in each temporal delay condition. First, we examined the condition without temporal delay ($t = 0$ msec) and found that causal ratings were significantly impacted by velocity ratio, $F(4,25) = 4.33$, $p < .01$, which replicated Michotte's original finding that causal perception of the launching effect depends on the ratio between the two objects' pre- and post-collision speeds. However, when a noticeable temporal delay was introduced, participants rated their causal impression primarily based on the length of the temporal delay—with much less attention given to velocity ratio—as there was no significant simple main effect of velocity ratio in the $t = 70, 140, 210, 280$ msec temporal delay conditions, $F(4,25)$ = 2.73, .58, 1.07, 1.36; $p$ = .052, .68, .39, .28, respectively.

The impact of relative vs. absolute motion on causal ratings of observed launching events was examined in the absence of temporal delay ($t$ = 0 msec), because it was in this condition that velocity had an impact on causal ratings. We found that causal ratings, in fact, were impacted by background movement at a 0 msec delay, $F(2,27) = 4.02$, $p = .03$ (see Figure 4, top left panel). Specifically, ratings in the rightward background condition were significantly smaller than ratings in the rest background condition, $F(1,28) = 7.94$, $p < .01$, as well as smaller than the leftward background condition, $F(1,28) = 5.35$, $p = .03$. These results indicate that when the relative motions of two colliding objects (with respect to a moving background) are slow, people are less likely to report that the motor object launches the projectile object forwards.

### Model Results

Predictions from the noisy Newton model with absolute velocity inputs are indicated by the black curves in the bottom panels of Figure 4. The model predictions were compared

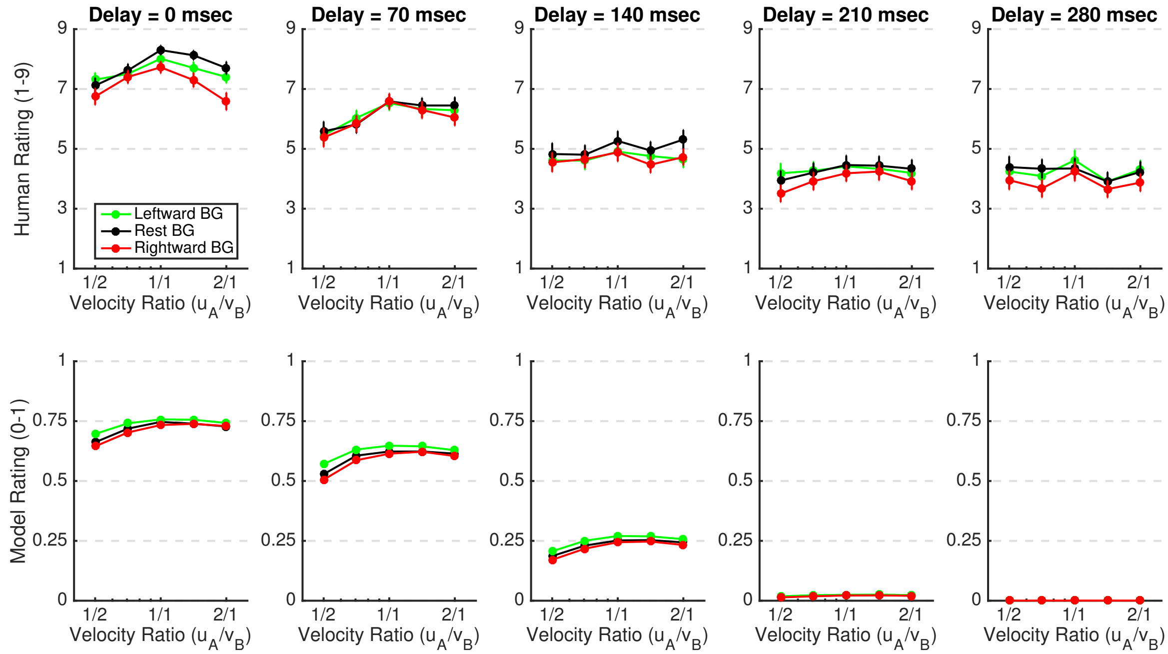


Figure 4: Human (top) and model predicted (bottom) causal ratings averaged across pre-collision motor speed ($u_A$) for each velocity ratio ($u_A/v_B$) condition. Separate plots (left to right) indicate different temporal delay conditions, and separate lines indicate different background (BG) movement conditions. Vertical error bars indicate standard error of the mean.

with human ratings in the $t$ = 0 msec condition because it was here that background movement had a significant impact. The absolute model reveals a fit of $r^2(13) = 0.314$ (95% CI = [0.004, 0.718]). The model's predictions are not influenced by the background movement condition, so it cannot explain the behavioral result that rightward background movement yielded lower causal ratings than resting and leftward background movement.

However, the causal noisy Newton model with relative motion inputs—as defined in the previous section—can qualitatively explain the impact of background movement on causal perception. Although the model fit with relative velocity inputs is comparable to the absolute velocity model fit, $r^2(73)$ = 0.376 (95% CI = [0.013, 0.761]), the predicted causal ratings are systematically influenced by background movement (see separate curves in bottom panels of Figure 4). While there was no observable difference between the leftward and rest background predictions, the model shows comparatively lower ratings in the rightward background condition compared with the leftward condition, which is consistent with human results. The model also captures Sanborn et al.'s (2013) finding that a 1/1 velocity ratio corresponds with peak causal ratings. Also note that model ratings achieve floor values at a temporal delay of approximately 210 msec. This occurs because large temporal delays disagree with the prior expectation of a 0 msec delay and thus generate a small temporal likelihood term. Human ratings also appear to reach floor values at around the same temporal delay. We chose the following model parameters in Experiment 2 to account for the human data: $\sigma$ = 972, $k_v$ = .319, $k_t$ = .004, $w_v$ = .059, $P(O|NC) = 2 \times 10^{-5}$. Note that a separate set of model parameters were chosen since the range of velocity input values was significantly greater than in the previous experiment.

## Discussion

The results reported herein demonstrate that (1) the motor object bias in mass ratio judgment is strengthened or attenuated when the background on which colliding objects travel moves either against or with the direction of their motion, respectively; and (2) impressions of launching are similarly influenced by moving backgrounds when there is no temporal delay between the movements of the two objects. The noisy Newton model (Sanborn et al., 2013) for object collisions was implemented and compared with human data in both a mass ratio judgment and causal rating task. For mass ratio judgment, the model with relative motion inputs accounts for human performance well across a range of experimental conditions. The goodness of the fit suggests that humans use perceived relative motion as the input to high-level cognitive systems when inferring observable physical properties. For causal perception, the model with relative motion inputs explains the finding that background movement with the motion of colliding objects negatively impacts causal ratings.

In summary, our results show that impressions about the attributes of—and relationships between—entities in the world are systematically influenced by low-level spatiotemporal characteristics in observed scenes. However, it remains unclear whether the influence of more abstract contextual properties (e.g., Mayrhofer & Waldmann, 2014) on causal impressions can be explained by their spatiotemporal characteristics alone. Another question is whether human counterfactual reasoning (Gerstenberg, Goodman, Lagnado, & Tenenbaum, 2015; Gerstenberg, Peterson, Goodman, Lagnado, & Tenenbaum, 2017) in object collision tasks are also systematically influenced by background movement. It would be beneficial for future work to explore these possibilities.

## Acknowledgements

The current work was supported by a National Science Foundation (NSF) Graduate Research Fellowship, University of California, Los Angeles (UCLA) Dissertation Year Fellowship, and NSF grant BSC-1655300.

## Appendix: Noisy Newton Model Details

The noisy Newton model for mass judgment uses Bayes' rule to calculate a posterior distribution of collision attributes, $A$, given noisy observable information $O$:

$$P(A|O) = \frac{P(O|A)P(A)}{P(O)}, \tag{1}$$

where $P(A)$ represents prior knowledge people have about hidden attributes in collision events. Those attributes are mass, $m_A$ and $m_B$, and restitution, $e$, which is a constant between 0 and 1 that represents the elasticity in a collision. The model assumes that all restitution values are equally likely and objects are more likely to be light than heavy: $e \sim$ Uniform(0,1); $m_A$, $m_B \sim$ Exponential(1). The $P(O|A)$ term indicates the likelihood of observed velocities ($O = u_A$, $u_B$, $v_A$, $v_B$) given a potential set of attributes ($A = e$, $m_A$, $m_B$). Here, $u_A$ and $u_B$ are the pre-collision velocities of Objects $A$ and $B$, and $v_A$ and $v_B$ are the post-collision velocities. Post-collision velocities are calculated based on the pre-collision velocities, the object masses, and the collision's restitution coefficient via the following equations:

$$v_A = \frac{m_A u_A + m_B(u_B + e(u_B - u_A))}{m_A + m_B} \tag{2}$$

$$v_a = \frac{m_B u_B + m_A(u_A + e(u_A - u_B))}{m_A + m_B} \tag{3}$$

A noisy observation model then links true, hidden variables $\bar{O}$ with observed variables $O$ such that their difference $\varepsilon$ is normally distributed in logarithmic space: $\varepsilon \sim \text{Gaussian}(0, k_x^2)$. Given a weighted logarithmic transformation function $f(x) = \text{sign}(x)\log(w|x|+1)$, the difference between observed and true observations is expressed as $\varepsilon = f(O) - f(\bar{O})$.

With the noisy observation model, the $P(O|A)$ in Equation 1 can be expanded to include both $O$ and $\bar{O}$:

$$P(O|A) = \int_{\bar{O}'} P(O|\bar{O}')P(\bar{O}'|A)\mathrm{d}\bar{O}', \tag{4}$$

where the $P(\bar{O}|A)$ term is further separated into initial and final velocities: i.e., $P(\bar{O}|A) = P(\bar{v}_A, \bar{v}_B|A)P(\bar{u}_A, \bar{u}_B)$. Note that pre-collision velocity does not depend on the collision attributes. Instead, values for $\bar{u}_A$ and $\bar{u}_B$ are drawn from the slow motion prior such that $\bar{u}_A$, $\bar{u}_B \sim$ Gaussian(0, $\sigma^2$). Post-collision velocities are then calculated from Equations 2 and 3.

The noisy Newton model can also be used to predict the marginal probability of a causal relationship, $C$, given noisy observable information, $O$:

$$P(C|O) = \frac{P(O|C)P(C)}{P(O|C)P(C) + P(O|NC)P(NC)}. \tag{5}$$

The $P(O|C)$ term in Equation 5 can be expanded to the following integral:

$$P(O|C) = \int_{\bar{O}',A'} P(O|\bar{O}')P(\bar{O}'|A',C)P(A')\mathrm{d}\bar{O}'\mathrm{d}A'. \tag{6}$$

Note that temporal delay, $t$, can be included as an observable variable with log-normal uncertainty and a delta function prior centered at 0 msec: $P(\bar{t}) = \delta(\bar{t})$. For the $P(O|NC)$ term, Sanborn et al. (2013) set this value as a free parameter in their model. The authors also made the assumption that causal and noncausal models were equally likely: i.e., $P(C) = P(NC)$.

## References

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.

Battaglia, P. W., Pascanu, R., Lai, M., & Rezende, D. J. (2016). Interaction networks for learning about objects, relations and physics. In *Advances in neural information processing systems* (p. 4502-4510).

Chang, M. B., Ullman, T., Torralba, A., & Tenenbaum, J. B. (2016). A compositional object-based approach to learning physical dynamics. In *Proceedings of the 5th international conference on learning representations* (p. 1-15).

Cohen, A. L. (2006). Contributions of invariants, heuristics, and exemplars to the visual perception of relative mass. *Journal of Experimental Psychology: Human Perception and Performance*, *32*(3), 574-598.

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.

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.*

Gerstenberg, T., Peterson, M. F., Goodman, N. D., Lagnado, D. A., & Tenenbaum, J. B. (2017). Eye-tracking causality. *Psychological Science*, *28*(12), 1731-1744.

Gibson, J. J. (1966). *The senses considered as perceptual systems*. Boston MA: Houghton Mifflin.

Gilden, D. L., & Proffitt, D. R. (1994). Heuristic judgment of mass ratio in two-body collisions. *Perception & Psychophysics*, *56*(6), 708–720.

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.*

Kubricht, J. R., Holyoak, K. J., & Lu, H. (2017). Intuitive physics: Current research and controversies. *Trends in Cognitive Sciences*, *21*(10), 749-759.

Leslie, A. M. (1982). The perception of causality in infants. *Perception*, *11*(2), 173-186.

Mayrhofer, R., & Waldmann, M. R. (2014). Indicators of causal agency in physical interactions: The role of prior context. *Cognition*, *132*(3), 485-490.

Michotte, A. (1963). *The perception of causality*. New York, NY: Basic Books.

Natsoulas, T. (1961). Principles of momentum and kinetic energy in the perception of causality. *The American Journal of Psychology*, *74*(3), 394-402.

Runeson, S. (1983). *On visual perception of dynamic events*. Stockholm, Sweden: Almqvist Wicksell.

Runeson, S., Juslin, P., & Olsson, H. (2000). Visual perception of dynamic properties: cue heuristics versus direct-perceptual competence. *Psychological Review*, *107*(3), 525–555.

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.

Saxe, R., & Carey, S. (2006). The perception of causality in infancy. *Acta Psychologica*, *123*(1), 144-165.

Schlottmann, A., & Anderson, N. H. (1993). An information integration approach to phenomenal causality. *Memory Cognition*, *21*(6), 785-801.

Scholl, B. J., & Nakayama, K. (2002). Contextual effects on the perception of collision events. *Psychological Science*, *13*(6), 493-498.

Smeets, J. B., & Brenner, E. (1994). The difference between the perception of absolute and relative motion: A reaction time study. *Vision research*, *34*(2), 191–195.

Todd, J. T., & Warren Jr., W. H. (1982). Visual perception of relative mass in dynamics events. *Perception*, *11*(3), 325-335.

White, P. A. (2006). The causal asymmetry. *Psychological Review*, *113*(1), 132-147.