Looking forwards and backwards: Similarities and differences in prediction and retrodiction

Kevin A Smith (k2smith@ucsd.edu) and Edward Vul (evul@ucsd.edu)
University of California, San Diego
Department of Psychology, 9500 Gilman Dr.
La Jolla, CA 92093 USA

Abstract

People must often infer what might have transpired in the past to bring about the present state of the world, a task called *retrodiction*. We hypothesize that retrodiction relies on similar cognitive mechanisms to *prediction* – inferring possible futures based on the present state of the world. Here we investigate how people perform on physical reasoning tasks that differ only in that people are asked to do either prediction or retrodiction. We find that average behavior is similar between tasks across a range of difficulty, though there was greater variability in retrodiction responses. We propose two ways in which prediction and retrodiction might be related; however, neither sufficiently explains the similarities and differences across tasks. We suggest that both tasks rely on similar cognitive processes, but that further research is needed to determine the exact relation.

Keywords: Prediction; Retrodiction; Physical Reasoning; Noisy Newton Physics

Introduction

People often predict what might happen next in the world to guide their actions: if someone throws a rock at you, where is it going and how do you move to avoid it? But in many cases, it is important to infer what happened in the past to bring about the present: where did the rock that whizzed by your head come from? These two types of tasks – prediction and retrodiction respectively – both require using information about the present to infer possibilities about a different period, but differ in the direction of time. Might these tasks therefore rely on the same cognitive mechanisms? And how does the difference in the direction of time affect how people perform these tasks?

While there has been relatively little research into retrodiction, there has been significant attention paid to prediction, with many arguing that we are constantly and automatically predicting the future in order to plan our actions (Bar, 2007; Buckner & Carroll, 2007; Seligman, Railton, Baumeister, & Sripada, 2013). Underlying these predictions are mental models that simulate future states of the world based on how we expect the world to unfold alongside our own actions (Grush, 2004).

Research focused on elaborative prediction (prospection) has suggested that this task of looking into the future relies on many of the same cognitive mechanisms as remembering the past (Schacter, Addis, & Buckner, 2007; Schacter et al., 2012). We similarly suggest that retrodiction uses the same cognitive mechanisms that allow us to extrapolate the world forward.

We propose two candidates for how retrodiction might employ the same extrapolation mechanisms as prediction. One candidate is that retrodiction is *reverse prediction* – just as we have mental processes to run the world forward, we might have complimentary processes to run the world backwards. Alternately, retrodiction might be *inverse prediction*. Here we draw parallels to the theory that vision is "inverse graphics" – we use a model of optics to infer how configurations of objects might give rise to various percepts, then condition on visual observations to determine what we are seeing (Kersten, Mamassian, & Yuille, 2004; Mansinghka, Kulkarni, Perov, & Tenenbaum, 2013). Retrodiction might involve positing potential past states of the world, then extrapolate the world forward in order to determine which prior configuration is most likely to have given rise to the current state of the world.

Here we focus on physical prediction and retrodiction for two reasons. First, we have evidence for how people run the world forward when engaged in physical reasoning. The noisy Newton theory suggests that people use unbiased physical dynamics to simulate how objects might move, but uncertainty about the latent state of the world might give rise to mis-predictions and even systematic errors (Battaglia, Hamrick, & Tenenbaum, 2013; Sanborn, Mansinghka, & Griffiths, 2013; Smith, Battaglia, & Vul, 2013; Smith & Vul, 2013). Second, physics is reversible in time, so if retrodiction is simply reverse prediction, we can make situations in which there should be no difference between how people do prediction and retrodiction. On the other hand, stochastic noise is not reversible in time, so if retrodiction is inverse prediction, we might observe different errors even with matched dynamics.

We first present an experiment in which we ask participants to make predictions and retrodictions with matched dynamics to find similarities and differences irrespective of discrepancies in the way the world unfolds forward or backward in time. We then describe two models that implement reverse prediction and inverse prediction and compare them to participants' retrodictions. Participants' average predictions and retrodictions were similar across a range of difficulties, suggesting commonalities in prediction and retrodiction; however, participants were more variable in their responses during retrodiction. We could explain prediction using the model of Smith and Vul (2013), but models of reverse and inverse prediction based on this forward model could not explain retrodiction, suggesting that despite being based on similar cognitive processes, there is a complex relationship between prediction and retrodiction. Future research can build on this result to further specify how prediction and retrodiction relate.

Experiment

Participants made judgments about the trajectory of a ball bouncing around a table simulated on a computer in two conditions. In the *prediction* condition, participants watched a ball in motion and indicated where it would cross a line if it continued along its trajectory. In the *retrodiction* condition, participants watched a similar ball in motion and indicated where on a line it must have come from so that it would follow the observed trajectory. Crucially, trials were matched across the two conditions such that where a ball would cross the line on a prediction trial was where the ball came from on the matched retrodiction trial.

Methods

Fifty UC San Diego undergraduates participated in this experiment for course credit.[1] All participants had normal or corrected to normal vision.

Participants viewed a computer monitor from a distance of approximately 60cm, which showed a "table" with dimensions of 1200x900px from a top-down view. On all trials, participants would watch a ball travel on the table for 750ms. After viewing this motion once, a vertical line would appear on the table, and the ball's motion would continuously repeat. In the *prediction* block, participants indicated where they believed the ball would first cross the vertical line if it continued its trajectory. In the *retrodiction* block, participants indicated where on the vertical line they believed the ball last passed before its observed trajectory. Participants indicated their response by clicking on the vertical line, after which point the motion animation loop would stop and participants would be provided feedback with either the motion of the ball continuing to the line (prediction), or the motion of the ball starting at the vertical line and continuing through the observed trajectory (retrodiction). Participants would earn 'points' based on how close to reality their response was, but these points were simply for motivation and were not used for any rewards or analysis. On each trial, we recorded where the participant indicated the ball would/did cross the vertical line and the time it took them to make that response from when the vertical line first appeared (see Figure 1).

Both the prediction and retrodiction blocks contained 150 trials each, matched across the blocks. On matched trials, the vertical line would be positioned at the same horizontal location, and the observed trajectory of the ball would be mirrored in time; in this way the correct response to where the ball would/did cross the line was the same in both the prediction and retrodiction blocks – we call this position the *actual crossing*. Trial paths were constrained such that the correct response was never in the top or bottom 50px (5.6%) of the table.

[1] One participant was excluded because she failed to follow directions, instead consistently responding as quickly as possible.

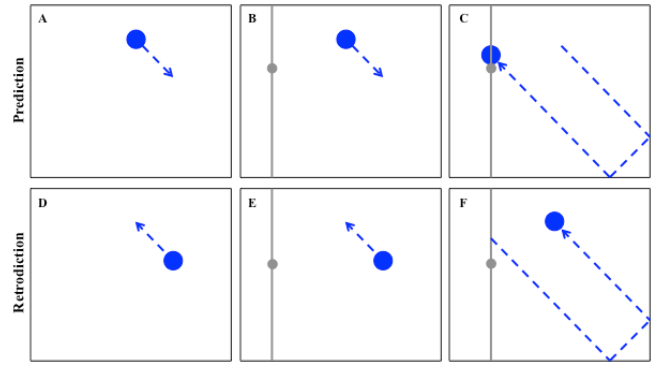

Figure 1: Diagram of a trial – prediction (*top*) and retrodiction (*bottom*). (A) In the prediction condition, participants would see a ball in motion. (B) After 750ms participants were asked to indicate on a line where the ball would cross; the ball would repeat its motion in a loop until a prediction was made. (C) After a prediction was indicated, the actual trajectory of the ball to the line was shown. (D) In the matched retrodiction trial, the mirrored ball motion was shown for 750ms. (E) Participants were asked where the ball last crossed a line. (F) Feedback was provided by showing the ball path from the line to the observations.

Each of the 150 trial pairs was categorized into one of six difficulty conditions, crossing the distance that the ball would need to travel to or from the vertical line (Short: 750px; Long: 1000px) with the number of the bounces that the ball would take on its unobserved path (0, 1, or 2). Trials were equally balanced into 25 from each category. The ball traveled at a constant speed of 500px/s for all trials.

Before each block, participants were given three practice trials to familiarize themselves with the task. Block order was randomized across participants, and all participants were given the same 150 matched trials in an order randomized for each block.

The trajectory of the ball both while observed and unobserved was determined by Newtonian physics, as implemented in the Chipmunk 2-D physics engine (Lembcke, 2011).

Results

Average responses by trial Participants responses for each trial were very consistent with one another (average split half correlation – prediction: $r = 0.97$; retrodiction: $r = 0.93$). We therefore aggregated participants' responses for each trial to determine where, on average, people believed the ball would go to or had come from on that trial.

Both the average prediction ($r = 0.87$) and retrodiction ($r = 0.85$) responses were correlated with the actual crossing location, suggesting that participants were taking into account trial differences (see Figure 2). Similar to Smith and Vul (2013), participants responded closer to the center of the screen than the actual crossing, and the center-bias and trial-by-trial variability increased as the distance and number of bounces increased.

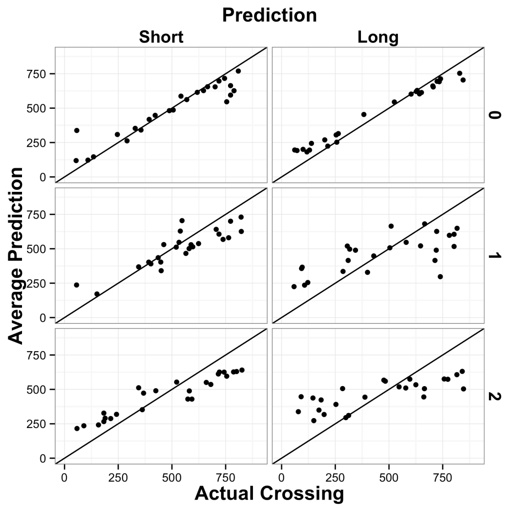

Figure 2: Average predictions (*left*) and retrodictions (*right*) versus where the ball would/did cross the vertical line, split by trial condition. Each point represents a separate trial. Error and trial-by-trial variability increases with difficulty, but differs only slightly from prediction to retrodiction.

On the other hand, predictions and retrodictions on matched trials were more correlated with one another ($r = 0.96$), and this correlation did not change appreciably with difficulty condition (from a minimum of $r = 0.94$ in the short, one bounce condition to a maximum of $r = 0.98$ in the long, no bounce condition; see Table 1). Because responses between the blocks are more related than each block's relation to ground truth (indeed, comparable to the internal consistency of each block), and this does not change with trial difficulty, it is likely that prediction and retrodiction rely on related extrapolation processes.

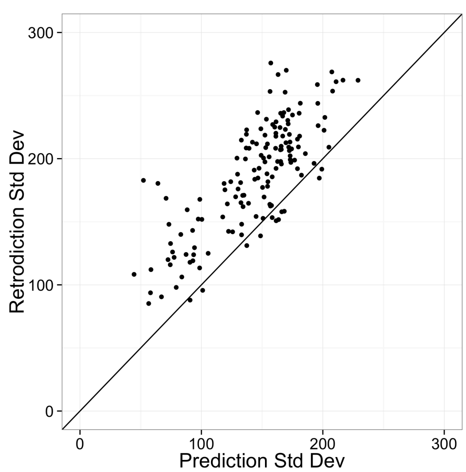

Figure 3: Variability in prediction versus retrodiction responses by matched trial. Each point represents a separate matched trial. There is general consistency in which trials are more or less variable, but retrodictions are more variable.

Variability of responses Although participants' average responses were very similar between the prediction and retrodiction, individual responses were more variable in the retrodiction block. We measured variability as the standard deviation of participants' responses within each trial. By this measure, trials that were variable in the prediction condition were also variable in the retrodiction condition ($r = 0.78$), but retrodiction trials were consistently more variable than prediction trials (see Figure 3).

Table 1: Prediction vs. retrodiction comparison. (1) By-trial correlation between prediction/retrodiction responses. (2) Median prediction response time (ms). (3) Median retrodiction response time (ms)

Trial condition	P/R cor	RT (Pred)	RT (Retro)
Overall	0.963	2,371	2,467
Short distance			
0 bounce	0.978	1,952	2,267
1 bounce	0.941	2,415	2,447
2 bounces	0.966	2,664	2,675
Long distance			
0 bounce	0.981	2,084	2,244
1 bounce	0.969	2,442	2,498
2 bounces	0.954	2,736	2,792

Reaction times In this experiment, participants were not given any time constraints – they could take as long as needed to make predictions or retrodictions. Thus differences between the two blocks could be masked if participants took significantly longer on one of the two blocks. There was evidence that participants took longer to do the retrodiction task than the prediction task (Wilcoxon rank test: $p = 0.021$), but this appears to be driven by slower responses on trials without unobserved bounces (see Table 1). Because we find similar amounts of center-shifting but increased retrodiction variability even when participants spent similar amounts of time across tasks, it is unlikely that our observations are biased by participants spending more time on the more difficult task.

Learning over time We assumed that participants would bring prior knowledge about physics to bear on the tasks we gave them in this experiment. We therefore provided feedback after every trial to encourage participants to be accurate. However, it is possible to learn physical contingencies from this feedback, and indeed we did find mild evidence for learning: by correlating absolute error (unsigned response less actual crossing) with trial order, we find a trend towards anticorrelation in the prediction task ($r = -0.020$, 95% CI = [-0.043, 0.002]) and modest anticorrelation in the retrodiction task ($r = -0.053$, 95% CI = [-0.075, -0.030]). Because of the small effect of learning and random ordering of trials, we do not believe that learning influenced any of the other results we found.

Reverse vs. inverse prediction

The behavioral results suggest that prediction and retrodiction are accomplished through similar mechanisms. Here we begin to investigate how they are tied together. We use the forward physical model of Smith and Vul (2013) to describe participants' predictions, then model retrodiction as both reverse and inverse prediction based on that framework.

Physical prediction model

In Smith and Vul (2013) participants were required to predict the motion of a ball on a computerized table, similar to the prediction task in this experiment. We found that participants' predictions could be well described by a noisy Newton model that assumed people had two classifications of uncertainty: *perceptual* uncertainty, which incorporated noise in judging the exact location and direction of movement of the ball immediately before the prediction, and *dynamic* uncertainty which accounted for noise that accumulated as the ball moved or bounced off of the sides of the table. In addition, we found that people were influenced by a prior belief that the ball would end up in the center of the table.

We fit the prediction responses in this experiment with the same model in order to quantify the levels of uncertainty our participants had in their forward models for this task. This forward physics model described participants' average responses by trial ($r = 0.95$; Figure 4, upper left) and standard deviation of participants responses within a trial ($r = 0.78$; Figure 4, lower left). This model assumed somewhat less uncertainty than the model of Smith and Vul (2013) to capture the lower variability in participants' responses (possibly due to providing unlimited time to make predictions). With a model of how participants simulated the world forward, we then investigated how this model might be used in retrodiction.

Retrodiction as reverse prediction

We considered two cases of reverse prediction. In the basic case, there is no difference between simulating the world forward or backward, and therefore there should be no differences between model fits for prediction and retrodiction for each trial. In the noisy retrodiction case, we assumed that the extrapolation dynamics would be the same as prediction, but considered that people might have difficulty reversing the motion of the ball; to model this we used the same dynamic uncertainty parameters from the forward model and re-fit perceptual uncertainty parameters to the retrodiction data.

Since the average retrodictions are highly correlated with average predictions, the basic reverse prediction model does describe the average by-trial retrodictions well, albeit more noisily ($r = 0.92$; see Figure 4, upper middle-left). Reverse model predictions also corresponded well with empirical retrodictions: we should expect a 1:1 relationship between participants' retrodictions and the model on average, and we found a slope of 0.99 (95% CI: [0.93,1.07]).[2] However, the model also predicted a significantly lower level of variability in peoples' responses than was actually observed (see Figure 4, lower middle-left).

The noisy reverse prediction model was also well correlated with participants' responses ($r = 0.92$) and produced less biased amounts of variability by trial (see Figure 4, middle-right). However, unlike the basic reverse prediction model, its predictions are biased as compared to participants' retrodictions: the slope against empirical data was 1.35 (95% CI: [1.27, 1.44]), suggesting that additional perceptual noise cannot account for the increased variability without producing model bias.

Because basic reverse prediction cannot capture empirical variability and noisy reverse prediction is biased, retrodiction must not simply be reverse prediction, but must require additional mechanisms to add variability without bias.

Retrodiction as inverse prediction

Inverse prediction involves proposing possible starting conditions then running the world forward to test how likely each starting condition is to have given rise to observations. As such, it is a candidate process that might give rise to greater variability in responses than reverse prediction. However, it requires defining both how initial conditions might be sampled, and how people determine whether a forward sampled ball path is likely or unlikely to have given rise to observations.

In the forward model, we assumed that people have a prior expectation on where the ball would cross the vertical line. If people share expectations across the prediction and retrodiction tasks, then this should be the prior expectation on the position the ball would start from – $p(y)$.[3] For the initial direction of the ball's motion – $p(\theta)$ – we assumed a relatively uninformative prior: since the vertical line was always to the left of the observed trajectory, we assumed that the ball must have some component of rightward motion, but that the direction was uniformly sampled from the 180° range of potential directions.

---

[2] Because the model predictions were estimated by simulation and therefore not exact, we calculated the slope with total least squares regression rather than ordinary least squares (Markovsky & Van Huffel, 2007).

[3] While we considered a uniform prior on starting position, this is logically inconsistent with the forward model and produced biased retrodictions; we therefore did not include these results.

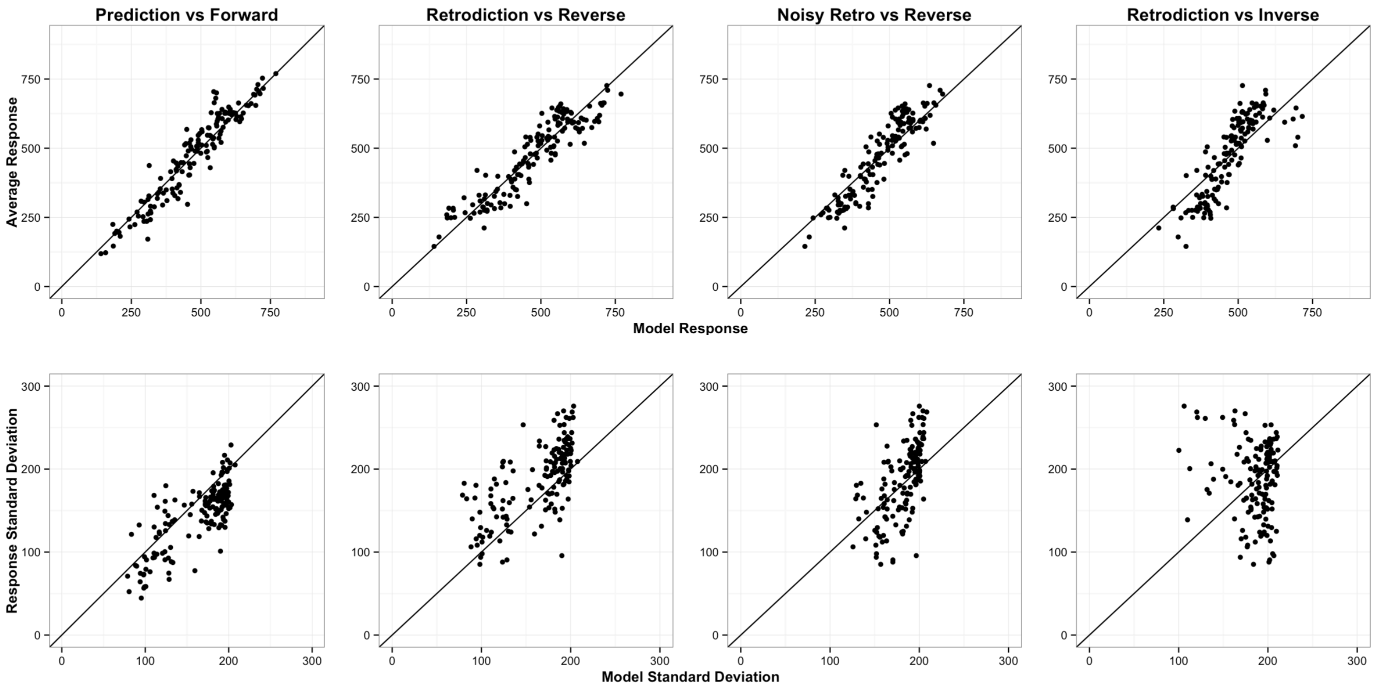

Figure 4: Correlation between model predictions (x-axis) and participants' behavior (y-axis). *Top:* Average responses by trial. *Bottom:* Standard deviation of responses by trial. *Left to right:* Forward model, basic reverse prediction model, noisy reverse prediction model, and inverse prediction model.

We assumed people would choose to accept a given starting position based on how well a simulated path emanating from that position matched the observed trajectory. Deviation for a single path was defined as the minimum summed squares deviation of the position of the ball along that path from the observed positions, sampled each 25ms (the time between observed frames). The probability that an observation came from a specific path was calculated as proportional to the negative exponentiation of this deviation, multiplied by an adjustment factor α – the one additional free parameter in this model; this formulation is equivalent to assuming that the path is correct, but there is isotropic Gaussian error in observations around that path, then asking what is the probability of those observations arising. To calculate the probability of getting a set of observations if the ball was initialized with a given y and θ, we sampled 50 paths using those starting parameters and the dynamic uncertainty fit from the forward model, then averaged the path probabilities. Finally, according to Bayes rule, the probability of selecting a given ending location was determined by adjusting the probability of getting the observations ($\boldsymbol{v}$) from a given value of y and θ, multiplying by the priors, and marginalizing over all possible starting values of θ:[4]

$$p(y|\boldsymbol{v}) \propto \int_{\theta} p\,(\boldsymbol{v}|y,\theta)p(y)p(\theta)d\theta$$

However, this formulation of inverse prediction did a poor job describing how people made retrodictions in aggregate. The model predictions of average response by trial were less well correlated with where participants actually indicated the ball came from (r = 0.83), and the model overcorrected for center-bias, suggesting that people should guess closer to the middle of the screen than they in fact did (see Figure 4, upper right). In addition, while the inverse prediction model did suggest a higher level of variability in participants' responses than the reverse prediction model did, it increased variability indiscriminately, suggesting that people should be highly variable on most trials, including many trials in which they were not (see Figure 4, lower right). Thus it is unlikely that people perform retrodiction by simply proposing candidate past world states and choosing one based on whether it might give rise to the present.

Discussion

In this experiment, we attempted to disambiguate the cognitive mechanisms underlying both prediction and retrodiction by giving people tasks matched in all dynamics except that one required prediction and the other retrodiction. We found that while there were center-biases in both prediction and retrodiction, participants' average

[4] Because these values could not be computed analytically, we used a grid search over possible values of y and θ in a 51x51 grid. We used LOESS smoothing with a low span parameter (0.02) to obtain probabilities at all points of y and θ, and to smooth out any variability due to the approximation in path sampling while keeping fine structure in the probability function.

predictions were the same as retrodictions on matched conditions, but responses were more variable on the retrodiction task. Participants' predictions could be well explained by a noisy Newton model of physics; however, computational models that treat retrodiction as reverse or inverse prediction fail to explain this result, instead assuming too much similarity to prediction (basic reverse), or too much difference (noisy reverse and inverse).

Table 2: Correlation between model average responses by trial and participants' average responses

Trial condition	Pred vs Forward	Retro vs Reverse	Noisy vs Rev	Retro vs Inverse
Overall	0.951	0.919	0.921	0.833
Short distance				
0 bounce	0.992	0.972	0.964	0.941
1 bounce	0.881	0.851	0.785	0.891
2 bounces	0.949	0.923	0.921	0.955
Long distance				
0 bounce	0.995	0.982	0.984	0.940
1 bounce	0.905	0.917	0.922	0.661
2 bounces	0.874	0.859	0.863	0.850

The similarity in behavior across the two tasks suggests that prediction and retrodiction do rely on similar cognitive mechanisms for extrapolation. However, the model results suggest that there is not a simple link between them – retrodiction is not just running the world backwards, nor is it naively sampling possible starting positions and running those forward until one starting position explains observations. This finding itself implies that prediction requires sophisticated cognitive mechanisms; if it were simply line extrapolation, then we would expect to find that retrodiction is reverse prediction – tracing the line backward. Our failure to explain this link may be due to the simplicity of the models studied. Perhaps retrodiction is inverse prediction, but people use a more sophisticated mechanism for sampling potential starting locations – for instance, if the ball was moving horizontally while observed, it will be less likely to have been traveling nearly vertically to begin. Or perhaps retrodiction relies on a mixture of forward and backward sampling to converge on a proposed past state of the world.

These findings suggest a parallel to the prospection literature: while imagining the future and remembering the past rely on the same constructive processes, there are differences between the two tasks in the precision of responses and difficulty to accomplish them (Schacter et al., 2012). Similarly, this experiment shows that prediction and retrodiction act in similar ways, but crucially that there is more variability in retrodiction. However, we find here that there is a complex relationship between the tasks of looking forward into the future and backward into the past. We hope that future research will help disentangle how various cognitive processes are shared and differ across these to tasks.

Acknowledgments

KS was funded by a UC San Diego Interdisciplinary Collaboratory Fellowship; EV was funded by National Science Foundation CPS Grant 1239323.

References

Bar, M. (2007). The proactive brain: Using analogies and association to generate predictions. *Trends in Cognitive Science, 11*(7), 280-289.

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

Buckner, R. L., & Carroll, D. C. (2007). Self-projection and the brain. *Trends in Cognitive Science, 11*(2), 49-57.

Grush, R. (2004). The emulation theory of representation: motor control, imagery, and perception. *Behavioral and Brain Sciences, 27*(3), 377-396.

Kersten, D., Mamassian, P., & Yuille, A. (2004). Object perception as Bayesian inference. *Annual Review of Psychology, 55*, 271-304.

Lembcke, S. (2011). Chipmunk 2D (Version 6.2.0): Howling Moon Software. Retrieved from http://chipmunk-physics.net/

Mansinghka, V. K., Kulkarni, T. D., Perov, Y. N., & Tenenbaum, J. B. (2013). *Approximate Bayesian image interpretation using generative probabilistic graphics programs.* Paper presented at the Advances in Neural Information Processing Systems.

Markovsky, I., & Van Huffel, S. (2007). Overview of total least-squares methods. *Signal processing, 87*(10), 2283-2302.

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.

Schacter, D. L., Addis, D. R., & Buckner, R. L. (2007). Remembering the past to imagine the future: The prospective brain. *Nature Reviews Neuroscience, 8*(9), 657-661.

Schacter, D. L., Addis, D. R., Hassabis, D., Martin, V. C., Spreng, R. N., & Szpunar, K. K. (2012). The future of memory: Remembering, imagining, and the brain. *Neuron, 76*(4), 677-694.

Seligman, M. E. P., Railton, P., Baumeister, R. F., & Sripada, C. (2013). Navigating into the future or driven by the past. *Perspectives on Psychological Science, 8*(2), 119-141.

Smith, K. A., Battaglia, P., & Vul, E. (2013). *Consistent physics underlying ballistic motion prediction.* Paper presented at the Proceedings of the 35th Annual Conference of the Cognitive Science Society, Berlin, Germany.

Smith, K. A., & Vul, E. (2013). Sources of uncertainty in intuitive physics. *Topics in Cognitive Science, 5*(1), 185-199.