# Learned prediction affects body perception

2 authors:

Benjamin J Balas
North Dakota State University
**139** PUBLICATIONS **2,191** CITATIONS

SEE PROFILE

Pawan Sinha
Massachusetts Institute of Technology
**203** PUBLICATIONS **7,292** CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Project Prakash View project

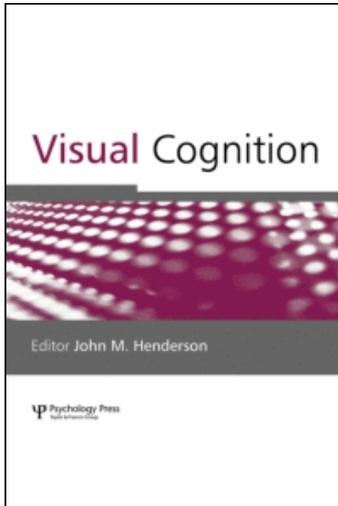Development of Face Perception after Sight Restoration (Project Prakash) View project

## Visual Cognition

## Learned prediction affects body perception

Benjamin Balas [a]; Pawan Sinha [a]

[a] Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA, USA

## PLEASE SCROLL DOWN FOR ARTICLE

# Learned prediction affects body perception

Benjamin Balas and Pawan Sinha

*Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA, USA*

Learning to recognize objects appears to depend critically on extended observation of appearance over time. Specifically, temporal association between dissimilar views of an object has been proposed as a tool for learning invariant representations for recognition. We examined heretofore untested aspects of the temporal association hypothesis using a familiar dynamic object, the human body. Specifically, we examined the role of appearance prediction (temporal asymmetry) in temporal association. In our task, observers performed a change detection task using upright and inverted images of a walking body either with or without previous exposure to a motion stimulus depicting an upright walker. Observers who were exposed to the dynamic stimulus were further divided into two groups dependent on whether the observed motion depicted forward or backward walking. We find that the effect of the motion stimulus on sensitivity is highly dependent on whether the observed motion is consistent with past experience.

Many intriguing results suggest that object motion and form both play an important role in object recognition, yet the relationship between spatial and spatiotemporal processing in object learning and recognition remains unclear. In the current study, our goal was to further characterize the nature of this relationship by examining how short and long-term exposure to a complex, moving object affects subsequent processing of static images of that object. Specifically, we asked how observed temporal asymmetries in object motion affect the ability to detect changes in object appearance.

The way an object moves is an important aspect of its appearance in many settings. Object motion provides information that can be used for

recognition independent of static form. For example, any individual image of a point-light walker does not evoke a very vivid percept of a human form, but once it begins to move the percept is irresistible (Johansson, 1973). Despite the sparsity of point-light stimuli, observers are capable of recovering the gender and identity of walking figures with good accuracy (Cutting, 1987; Dittrich, 1993; Kozlowski & Cutting, 1977; Pollick, Lestou, Ryu, & Cho, 2002). In other cases where static form is impoverished, motion features can support object or person identification (Knappmeyer, Thornton, & Bülthoff, 2003; Rosenblum et al., 2002). However, even under clear viewing conditions, object-centred and observer-centred motion can be important factors in recognition (Newell, Wallraven, & Huber, 2004). Object motion can thus be considered a useful set of additional measurements available for object recognition alongside static features.

These results suggest that object motion can play a vital role in object recognition, but at present there are few models that offer a concrete description of how object motion contributes to subsequent object perception and recognition. We continue by discussing two important proposals regarding the motion/form relationship: (1) "Temporal association" between temporally proximal images, and (2) prediction of future object appearance.

## TEMPORAL ASSOCIATION: TEMPORALLY SYMMETRIC GENERALIZATION VIA OBSERVED OBJECT MOTION

A recent proposal regarding the interaction of object motion with object form suggests that temporal proximity between images leads to the creation of a common representation for images presented close together in time. This proposal, which we shall refer to as the "temporal association" hypothesis, provides a method for learning invariant object recognition from natural image sequences. Many changes in viewing conditions preserve identity while drastically altering image-level features (Moses, Adini, & Ullman, 1994), making generic object recognition an extremely difficult problem. The temporal association hypothesis offers a principled means of linking different images to the same physical object; observing appearance change over time leads to generalization (and thus, invariance) over distinct images. Behaviourally, evidence for such linkages has been demonstrated for faces and novel objects using paradigms in which temporal association between images leads to increased confusability between objects presented close in time (Cox, Meier, Oertelt, & DiCarlo, 2005; Wallis, 1996; Wallis & Bülthoff, 2001). Several computational models of this process have been implemented as well, lending credence to the idea that learning to recognize a novel object under different viewing conditions may be facilitated by associative learning

over time (Foldiak, 1991; Stone, 1999; Stone & Harper, 1999; Ullman & Bart, 2004; Wallis, 1998).

The influence of associative learning is often revealed by demonstrating that an observer or a model produces the same response to previously distinct stimuli. The end result of associative learning in such a framework can thus be described as robust generalization over images that occur within some temporal interval. As a result, we consider such associative learning to be a temporally symmetric mechanism. Even though instantiations of associative models generally require a temporal "trace" that associates past stimuli with present inputs, the end result is a common representation for those two stimuli that does not take image order into account. Indeed, in most examinations of the effect of temporal association on object recognition and memory, temporal asymmetries during learning are either nonexistent or irrelevant to the task.

## TEMPORALLY ASYMMETRIC PREDICTION VIA OBSERVED OBJECT MOTION

Temporally asymmetric events are of course very common in the natural visual world. There is also substantial evidence that observers use their knowledge of motion asymmetries in the natural world to predict future events, leading to measurable differences in change detection and apparent motion (Freyd, 1983). There has also been a great deal of work describing a phenomenon known as "representational momentum", in which observers automatically "advance" a halted motion sequence forward in time (Freyd & Finke, 1984). These, and other similar effects such as the "flash-lag" illusion (Nijhawan, 1994), suggest that observers constantly use temporal contingencies in the world to formulate predictions regarding future stimuli. Effects of image order on recognition are evident for very complex biological motion stimuli, including human bodies (Verfaille & Daems, 2002) and faces (Lander & Bruce, 2000). It has also been shown that static images of "frozen" motion can give rise to activation in cortical areas devoted to motion perception (Kourtzi & Kanwisher, 2000), demonstrating that the prediction of future appearance is both psychologically and neurally real.

Recently, there have also been several studies that indicate that moving objects are represented in terms of a directed view sequence, or "spatio-temporal signature" that incorporates the observed trajectory of appearance change for arbitrary objects. Observers who learn a set of novel objects undergoing rigid rotation are significantly impaired at recognizing these objects if the direction of rotation is reversed at test (Stone, 1998). The presence of object motion that is inconsistent with training appears to overshadow the preservation of static form in this case, as evidenced by

multiple studies with rigid and nonrigid objects (Chuang, Vuong, Thornton, & Bülthoff, 2006; Stone, 1999; Vuong & Tarr, 2004).

## BODY PERCEPTION AS A TOOL FOR STUDYING PREDICTION AND OBJECT PERCEPTION

In the current study, our goal was to complement recent work regarding the "temporal association" hypothesis with an investigation of the influence of temporal asymmetry on sensitivity to changes in object form over short and long timescales. To achieve this goal, we chose to examine the perception of the human body.

Human body perception is a valuable test case for any theory positing that experience with a dynamic object should influence subsequent performance. First, the human body moves nonrigidly, limiting the utility of purely static representations of 3-D form (Biederman, 1987; Marr & Nishihara, 1978). Second, observers have a great deal of visual experience with moving bodies. Knowledge of biomechanical constraints and the expected form of the human body can profoundly affect "low-level" processing. For example, Shiffrar and Freyd (1990) demonstrated that perceived apparent motion of the human body is determined both by display speed and observers' knowledge of possible and impossible motions. Similarly, knowledge of allowable body movements affects the priming of interpolated frames from apparent motion sequences (Kourtzi & Shiffrar, 1999). Top-down influences on the perception of body-like figures were also reported by Sinha and Poggio (1996), who demonstrated that a rigidly rotating wire "walker" was generally perceived as nonrigidly deforming, presumably because observers expected the human-like form to move in a manner consistent with locomotion. Finally, body perception is also useful in that visual experience with human bodies is temporally asymmetric; forward walking is observed far more often than backward walking. Taken together, all of these factors make the human body an ideal subject for investigating the influence of observed temporal asymmetries over short and long timescales.

In our experiments, we used a simple same/different task to measure observers' sensitivity to small changes in object form. First, we assessed the influence of learned prediction by characterizing observers' ability to detect differences between two images of a normal walking body arranged in natural or reversed temporal order. We carried out this procedure at multiple viewpoints and with inverted bodies to test the generality of any influence of learned temporal asymmetry. Second, we determined how recent exposure to a dynamic stimulus (a walking human) affected subsequent performance of this task. To do this, we exposed different groups of observers to either

forward or backward walking during a 10-minute passive training period to determine if the direction of temporally asymmetric change influenced sensitivity over a short timescale. In all cases, we characterized behaviour within a signal detection framework, allowing us to differentiate between changes in sensitivity and response bias resulting from our experimental manipulations. In both experiments, our main question was whether or not temporal asymmetry (the direction of walking observed either over the lifetime or the laboratory session) impacts an observer's ability to detect changes in object appearance. We began by asking how observers' lifetime experience with forward walking affects sensitivity to appearance change.

## EXPERIMENT 1: DOES LEARNED PREDICTION AFFECT STATIC FORM SENSITIVITY?

In our first task, we asked whether or not prior knowledge concerning expected movement of the human body has consequences for change detection. We measured the relative sensitivity to "forward" and "backward" image changes across three different views of the same walking figure in upright and inverted orientations. This allowed us to ask whether or not there are consistent effects of appearance prediction on sensitivity, and also whether any such effects are dependent on viewing angle or the familiarity of the stimulus.

### Method

*Subjects.* A total of 12 volunteers participated in this task, all of whom were naïve to the purpose of the experiment. All participants were between the ages of 18 and 35 and reported normal or corrected-to-normal visual acuity.

*Stimuli.* All images were created using Poser 6, a 3-D graphics tool for rendering and manipulating models of human bodies (Curious Labs). A male figure was created using the software's standard "walk generator" settings and rendered from three viewpoints (side, three-quarters, and frontal views) while walking in place (as on a treadmill) at a normal speed. Figure 1 contains example images of the model at each of the three rendered views. Sixty images were rendered at each view depicting the model carrying out a complete walking cycle of two steps, allowing us to continuously loop these images to display ongoing walking. Each static image was $278 \times 484$ pixels in size and contained 256 grey levels.

Creating images of a walking human using graphics software had several advantages. Rendering the same physical event from multiple viewpoints is easily accomplished in the Poser environment, and lighting conditions,

**Figure 1.**   Example views of the male walker created for the experiments. The images depict the model in the same pose for each of our three views.

walking speed, and background are also easy factors to control. Of course, these gains come at the expense of ensuring a "natural" input (a point we will raise again in the Discussion). For our purposes, the Poser images offered a convenient alternative to filming real-world walking sequences in which lighting, viewpoint, and background were all well-controlled across images.

*Procedure.*   All participants performed a change detection task using images of the walker. Stimuli were presented in blocks according to viewpoint. Block order was counterbalanced across subjects.

On each trial, observers saw two still images of the walker for 500 ms each, separated by a 500 ms interstimulus interval (Figure 2). Nonmatching stimulus pairs were always two frames apart in the original rendered sequence, but from trial to trial the position of the two frames within the larger sequence varied. This means that the amount of image change between stimulus pairs varies somewhat from trial to trial, but the spacing within the sequence is constant. At test, the second image was always translated relative to the first by $+/-10$ pixels horizontally and vertically (independent random selection for each direction on each trial). Observers were instructed to press "1" if they believed the two images were identical, and "0" if they believed they were different. Errors were indicated with an auditory stimulus.

Image pairs contained either two upright or two inverted images that could be identical, different in "forward" order, or different in "backward" order. Each "forward" pair contained two images that were in the correct
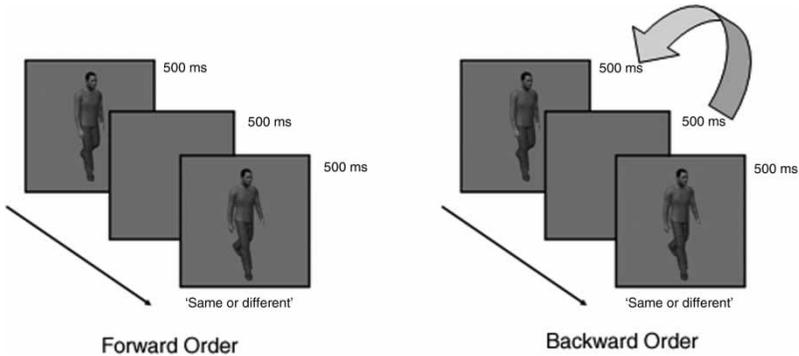
**Figure 2.** On each change detection trial, two images were presented sequentially for 500 ms each. Images could be identical, different, and in forward temporal order, or different and in reverse temporal order. Each forward pair had a corresponding backward pair created by swapping image order for presentation. Both upright and inverted pairs were presented during the same test session.

temporal order for forward walking, and corresponding "backward" pairs contained the same images in the opposite order. This manipulation guarantees that low-level differences between the two images are matched across the "forward" and "backward" conditions.

At test, the order of pair presentation within each block was randomized separately for each subject. Within a block, there were 42 upright "different" pairs (21 "forward" and 21 "backward") as well as 42 upright "same pairs". The same image pairs were inverted for a grand total of 168 trials per block.

All stimuli were presented on a 19-inch Sony monitor. Participants were seated approximately 50 cm from the monitor, such that all images of the walker subtended approximately $10 \times 5$ degrees of visual angle. Stimulus timing and response collection was monitored by software written using the Matlab Psychophysics Toolbox (Brainard, 1997; Pelli, 1997).

## Results

*Sensitivity to image change.*    This experiment allowed us to characterize the effect of a robustly observed temporal asymmetry (a lifetime of observing forward walking) on change detection between images placed in typical or atypical temporal order. Our primary question is whether or not there is any influence of frame order ("forward" or "backward") on this task. A secondary question is whether or not any observed effect generalizes to the different views used in this task, or to the inverted bodies. To answer these questions, we calculated $d'$ for both forward and backward image pairs of all views in the upright and inverted orientation. Graphs of these results are displayed in Figure 3.
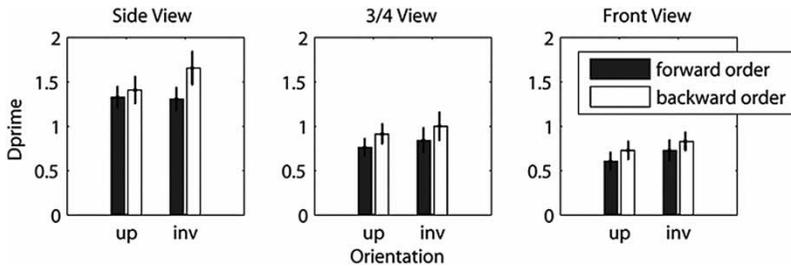
**Figure 3.** Discriminability for upright and inverted stimuli presented at three views in either forward or backward temporal order. We observe main effects of temporal order (favouring backward presentation) and view (favouring larger image differences) with no significant interactions or other main effects. Error bars represent $1+/-$SEM across the group data for each condition.

When calculating $d'$ for forward and backward pairs of the same view and orientation, the same false alarm rate was applied to both $d'$ calculations since there was only one kind of "same" trial for each combination of view and orientation. This means that differences in $d'$ between the forward and backward conditions within each view and orientation condition are actually only differences in hit rate. However, we report $d'$ values in all cases so that we may compare sensitivity across different view and orientation conditions, which did have distinct "same" trials and may thus have different false alarm rates.

A $3 \times 2 \times 2$ repeated-measures ANOVA with view, orientation, and temporal order as within-subjects factors yielded a main effect of view, $F(2, 10) = 17.93$, $p < .001$, partial $\eta^2 = .620$, and a main effect of order, $F(1, 11) = 9.43$, $p = .011$, partial $\eta^2 = .46$. No other main effects or interactions were significant.

*Response bias.*    To further compare performance in the upright and inverted conditions, we also report the response criterion for upright and inverted images as a function of view. This measure describes an observer's tendency to respond "same" or "different" irrespective of the actual stimuli presented during the experiment. This analysis allows us to determine whether or not the profound imbalance between prior exposure to upright and inverted bodies led to the application of different response criteria to upright and inverted images. A graph of these results by view and orientation condition is displayed in Figure 4.

A $3 \times 2$ repeated-measures ANOVA with view and orientation as within-subjects factors revealed only a significant effect of stimulus orientation, $F(1, 11) = 8.127$, $p = .016$, partial $\eta^2 = .425$. This result indicates that across all views, upright stimulus pairs evoked more "same" responses than
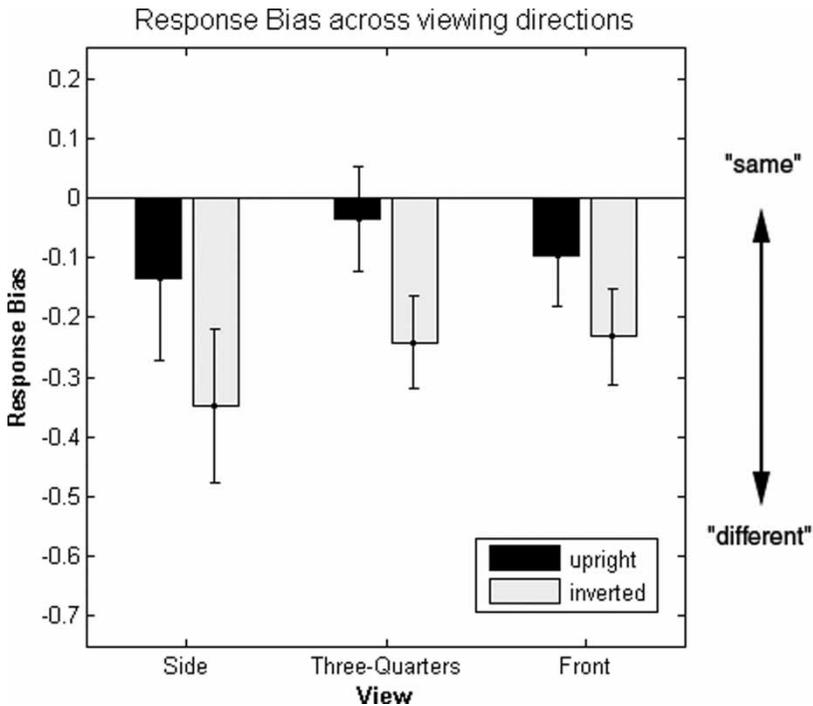
**Figure 4.**    Response bias for upright and inverted stimuli presented at three views in either forward or backward temporal order. We observe a main effect of orientation such that upright stimuli generate more "same" responses overall. Error bars represent $1+/-\mathrm{SEM}$ across the group data for each condition.

inverted pairs, despite the fact that there appears to be no difference in sensitivity between the upright and inverted pairs.

## Discussion

The data from Experiment 1 allowed us to answer our primary question of whether or not there is an influence of long-term temporal asymmetries on the ability to detect subtle changes in appearance. We found that "forward" differences are harder to detect than "backward" differences, even when the low-level stimulus differences between these conditions are perfectly matched. This suggests that temporal asymmetry does have an influence on the ability to detect change.

Why should "forward" differences be more difficult to detect than their reversed counterparts? One might think that familiarity with forward walking might lead to a strong hypothesis about the probability of

"forward" versus "backward" image changes, and it has been shown that prior knowledge about the probability of change in classic change detection tasks increases observers' abilities rather than impairs them (Beck, Angelone, & Levin, 2004). However, we suggest that our results are in accord with previous work implying the existence of "automatic prediction" in cases where an asymmetry has been learned (Freyd, 1983). Automatically "advancing" the first image in a "forward" pair makes it resemble the second image more closely, whereas the same operation performed on a "backward" pair increases the dissimilarity between the two stimuli. This is also directly in line with previous accounts of "representational momentum", with the caveat that our stimuli in Experiment 1 have no explicit "momentum" because there are only two test frames (rather than an extended sequence culminating in a single test image). Here, the observers' knowledge of how the object *should* move induces a direction and "momentum".

In terms of our second question, we also found that the effect of temporal order was evident at multiple views and invariant to picture-plane rotation. The lack of an interaction between view and temporal order suggests that prediction occurs within an object-centred frame of reference. This is somewhat at odds with characterizations of biological motion perception as strictly view dependent (Verfaille, 1993), insofar as prediction appears to take place here in a view-independent way. However, it is not clear that view dependence in recognition performance should imply view-dependent prediction. Our results complement existing work by demonstrating a robust and consistent prediction effect at multiple viewpoints. An interesting extension of this task would be to examine whether or not prediction can be transferred across a change in view (though judging whether a human body has moved slightly across a view change would likely be quite difficult for most observers).

Beyond the generality of our effect across multiple views, it is somewhat surprising to see that the $d'$ scores from the inverted stimuli were basically identical to those from upright images. The only difference we observed between upright and inverted pairs is a criterion shift (subjects seeing upright images are more likely overall to respond "same"), not a change in relative or absolute discriminability. Observers' heavily biased experience viewing upright, rather than inverted, bodies translated only into a difference in the chosen threshold for reporting image differences rather than in the ability to perceive them.

Overall, the data from Experiment 1 support the notion that temporal asymmetry exerts an influence on change detection at multiple object poses and orientations. We continued by examining whether or not exposure to a particular direction of appearance change can influence sensitivity over relatively short time scales.

## EXPERIMENT 2: THE EFFECT OF RECENT EXPERIENCE ON IMAGE DISCRIMINATION

In our second task, we further investigated the nature of object learning from dynamic stimuli by examining whether or not a temporal asymmetry observed in recent, brief exposure to a moving object affects the ability to detect subtle changes in object appearance. To do this, we repeated the change detection task employed in Experiment 1 with two new groups of observers who were first exposed to 10 minutes of either forward or backward walking. As before, we measured sensitivity to image change and response bias in each group of observers, for upright and inverted images of the human body.

### Method

*Subjects.* A total of 24 volunteers participated in this task. All participants were between the ages of 18 and 35 and reported normal or corrected-to-normal visual acuity. All participants were naïve to the purpose of the experiments and had not participated in Experiment 1.

*Stimuli*. The images of our walker performing the normal gait as described in Experiment 1 were also used here. However, we limited ourselves here to the three-quarter view only.

*Procedure.*   Volunteers who were placed in either "motion" group began by viewing the upright model walking either forward (Group 1) or backward (Group 2) for 10 minutes. "Backward" walking in this task means nothing more than time-reversed forward walking. A real human attempting to walking backwards would probably look quite different than our time-reversed movies, making this stimulus a poor approximation of whatever experience an observer might have with backward locomotion. However, the use of the reversed sequence provides a good control for low-level image structure and dynamical properties of the stimulus.

Viewing was broken up into ten 1-minute long blocks, during which the 60 images of the walker (three-quarter view) were continuously looped at a frame rate of 30 frames per second (with a 60 Hz refresh rate). To ensure that observers attended to the animation during each block, a "cue-dot" detection task was administered during presentation. During playback of the walking stimulus, a small red dot ($\sim 0.5$ degrees of visual angle in diameter) was drawn at a random location on the image at randomly selected times for a duration of 32 ms. Observers were instructed to press the spacebar as quickly as possible every time this dot appeared. All observers were highly accurate at reporting the presence of the dot, indicating that they

were indeed attending to the area of the screen subtended by the stimulus (though perhaps not attending directly to the walking stimulus itself). The fact that observers needed to rely on fairly diffuse attention over the visual field to carry out the secondary task also helps control for the fact that backward walking is in some sense "more interesting" than forward walking given its novelty.

Immediately following the exposure period, all observers carried out the change detection task described in Experiment 1 using the images of the three-quarter walker only. There were no other differences in design parameters or procedure.

## Results

*Sensitivity to image change.* We calculated $d'$ for forward and backward discrimination of upright and inverted stimuli as in Experiment 1, noting again the shared false alarm rate for stimuli of the same orientation. These results are summarized in Figure 5.

A $2 \times 2 \times 2$ mixed-design ANOVA was carried out, with temporal order and orientation as within-subject factors and forward versus backward training motion as a between-subjects factor. The only significant effect was a main effect of training motion on discriminability, $F(1, 22) = 7.54$, $p = .012$, partial $\eta^2 = .255$. No other main effects of interactions reached significance. Exposure to forward motion prior to performing the discrimination task resulted in significantly better performance than exposure to backward motion composed of the same frames played in reverse order. An important feature of the data, however, is that observers
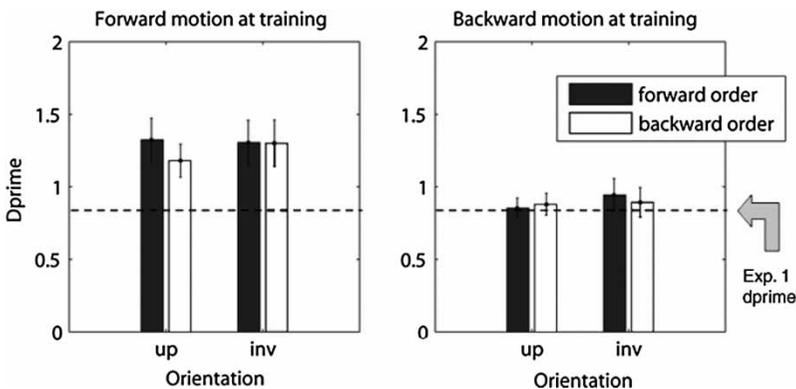


**Figure 5.** Discriminability for upright and inverted stimuli in forward or backward temporal order following exposure to a dynamic stimulus. Observing forward motion results in significantly better performance. Error bars represent $1 +/- SEM$ across the group data for each condition.

who watched backward motion were not "bad" at the task. Rather, forward motion appeared to have a facilitative effect (improving performance above the level observed in Experiment 1).

*Response bias.*   We continued by examining observers' response bias in the upright and inverted case in both training conditions. As previously, we discarded the distinction between "forward" and "backward" pairs in this analysis due to the shared false alarm rate within each orientation condition. In Figure 6, we display a graph of response bias in both groups alongside the data reported in Experiment 1 for the three-quarter condition. We can thus compare response bias given exposure to forward walking, exposure to backward walking, and also given no pretask exposure.

There was clearly no difference in the response criterion between the two motion-exposure conditions. However, it is evident from the figure
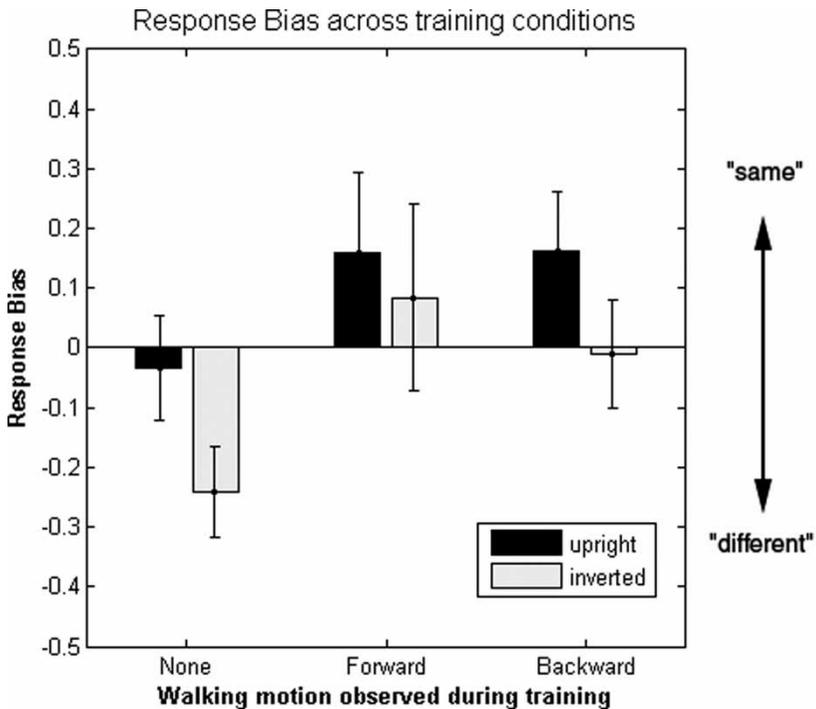


**Figure 6.**   Response bias following exposure to either a forward or backward walker (middle and right) or given no exposure prior to the change detection task (left). There is no difference at all between the two motion conditions, but there is a clear difference between receiving exposure and receiving none. Error bars represent $1 +/ -$ SEM across the group data for each condition.

that there is a difference between having received no exposure versus exposure to either dynamic stimulus. To confirm this effect, a $2 \times 2$ mixed-design ANOVA was carried out with orientation as a within-subject variable and exposure to motion (no exposure vs. any exposure) as a between-subjects variable. We found both a main effect of orientation, $F(1, 34) = 11.94$, $p = .001$, partial $\eta^2 = .260$, and a main effect of exposure to motion, $F(1, 34) = 4.91$, $p = .034$, partial $\eta^2 = .126$, confirming that upright pairs evoked more "same" responses and also that exposure to the moving walker also induced a bias to respond "same" more often.

## Discussion

The data from this task provides further support that temporal asymmetries in observed object motion influence subsequent object processing. Though the predictive effect observed in Experiment 1 appears to have been "washed out" by exposure to a moving walker prior to performing the change detection task, there is a strong effect of which sequence observers were exposed to. Under a purely associative mechanism (in which time has no "arrow"), there should be no difference between observing forward or backward locomotion during the exposure period since both sequences place the same images close together in time. Thus, we conclude that temporal asymmetry (the direction of image juxtaposition) can influence object perception after recent, and relatively brief, exposure to a moving object.

In terms of response bias, observers' response criterion "c" displayed the same "inversion effect" observed in Experiment 1. Upright images received more "same" responses than inverted images. Furthermore, the comparison between the criterion data from both experiments is interesting in that exposure to either forward or backward motion appears to elicit more "same" responses across both orientation conditions. We suggest that this is consistent with temporally symmetric association between images operating at the level of subjective bias. Exposure to the moving walker in both conditions seems to have led observers' to hold out for larger perceived differences between static images before producing a "different" response. However, the criterion difference between upright and inverted stimuli was preserved following exposure to the training stimulus, indicating that the position of the criterion for inverted stimuli is modulated by observed motion to roughly the same extent as the criterion for upright stimuli. This is an additional instance of observed object motion exerting a broad effect across manipulations of the original stimulus.

## GENERAL DISCUSSION

The results we have reported in Experiments 1 and 2 both provide evidence that temporal asymmetries in observed sequences have consequences for subsequent detection of image change. We close by discussing the results of both experiments in relation to one another, and suggest interesting topics for further inquiry.

Experiment 1 allowed us to test the effect of a learned bias in experience (observation of forward vs. backward walking) on discrimination for upright and inverted orientations of the human body. The observed difference in sensitivity favouring detection of "backward" image differences across views is consistent with automatic predictive mechanisms that project images forward in time. Furthermore, our sensitivity data ($d'$) demonstrate that this predictive process produces an effect of approximately equal magnitude at multiple viewpoints and is robust to picture-plane rotation of the learned stimulus. Our measurements of observers' response bias indicate that there is an effect of long-term exposure to upright images of a walker, but that this effect is evident in the criterion for reporting perceived appearance differences rather than in the ability to perceive those differences. Specifically, extensive experience with upright bodies leads to an elevated threshold for reporting appearance differences, which is broadly consistent with previous reports that observed temporal association contributes to an increased tendency to "bind together" previously distinct stimuli (Cox et al., 2005; Wallis & Bülthoff, 2001). These results by themselves are an interesting complement to previously reported results, suggesting a deeper complexity to the relationship between learned object motion and object processing.

The results of Experiment 2 are equally intriguing. First of all, exposure to either forward or backward walking erases the previously observed asymmetry in sensitivity for forward versus backward image pairs. Second, we find that exposure to forward walking leads to better sensitivity than exposure to backward walking for upright and inverted images. This difference between exposure to forward and backward motion is not predicted by an associative mechanism since temporal association in its simplest form only "knows" what pairs have appeared close together in time, not which image came first. In terms of subjective bias, the response criteria measured in this experiment appear to be "blind" to the direction of observed appearance change, but differ significantly from those measured in Experiment 1. In particular, recent experience with a moving, upright body leads to an elevated threshold for reporting appearance differences.

The results we report here are intriguing, but present several interesting questions. We address several of them next.

## Does the change detection task reveal aspects of static form representations, dynamic representations, or both?

An important issue to raise with regard to the "change detection" task used in both Experiments 1 and 2 is the extent to which the stimulus display parameters permit the perception of apparent motion between the sample and test frame of an image pair. The temporal parameters we selected (as well as the amount and unpredictability of spatial jitter between images within a pair) were all chosen in an attempt to minimize the percept of motion over the course of a trial; however, we cannot unequivocally state that we removed all motion from the stimuli. This leads to an important question that the current data does not allow us to resolve: Does temporal asymmetry affect "static" representations of form, "dynamic" representations of form, or both? This is a very difficult question for future studies to resolve, since fully removing all sense of motion from pairs of stimuli like those used in the current study would likely require fairly extreme measures that would certainly interfere with performance. That said, a comprehensive study of how learned temporal asymmetry affects the perception of "real" motion, apparent motion of varying "strength", and (if possible) purely static perception would be a valuable contribution. In terms of the current data, we have opted to discuss the results in terms of "change detection", rather than making an explicit claim as to whether it is the representation of static form that is changing or some representation of object dynamics.

## Why is there no effect of inversion on sensitivity in any of these tasks?

In both experiments, our measurements of sensitivity for upright and inverted bodies were almost identical. Observers were equally able (in terms of $d'$) to report image-level differences in a highly familiar orientation and a highly unfamiliar orientation. Furthermore, the magnitude of our observed effect of temporal asymmetry was also roughly equal across stimulus inversion. Why should this be the case?

Could the lack of an inversion effect be a byproduct of the overall familiarity of the human body? Previous reports of body-inversion effects (Pavlova & Sokolov, 2003; Reed, Stone, Bozova, & Tanaka, 2003) indicate that human body perception is not in general immune to inversion. Furthermore, inversion effects are often most pronounced for stimuli processed at an "expert" level (Diamond & Carey, 1986; Gauthier & Tarr, 1997; Yin, 1969). Indeed, recent results suggest that inversion effects emerge as expertise is acquired for arbitrary stimuli (Husk, Bennett, & Sekuler, 2007), which speaks strongly against the possibility that overexposure to

images of human bodies underlies this feature of the data. However, we have not only employed images which individuals are highly familiar with, but we have similarly considered only highly familiar patterns of motion. It is possible that observers learn orientation-invariant representations of real human motion, leading to robust change detection abilities. The use of an unfamiliar gait could inform this issue (though an unfamiliar gait also leads to unfamiliar images as well). Overall, using novel object classes would help determine the generality of the effects we have observed, which we must presently concede may be specific to either human body perception, or specific to the perception of forms with familiar motion.

A second possibility is that our digitally created stimuli may lack some aspects of natural human walking that could drive an inversion effect. In particular, the effects of gravity may not be accurately rendered in our images, leading to inexact and impoverished postural cues that might play a primary role in differentiating the perception of upright and inverted bodies. Although the Poser program is an impressive and thorough environment in which to create naturalistic stimuli, we must concede that we are ultimately subject to what limitations it has in modelling the complete set of physical factors that influence walking behaviour. The use of natural stimuli that reflect the postural effects of gravity on locomotion would thus be an interesting complement to the current results.

## Why might exposure to forward walking lead to increased sensitivity?

The sensitivity data from Experiment 2 demonstrates that exposure to familiar motion (forward walking) leads to better sensitivity than exposure to unfamiliar (backward) motion. This poses two challenges for a comprehensive theory. First, association between temporally proximal images generally leads to decreased, rather than increased sensitivity (Wallis & Bülthoff, 2001). Second, how does temporal asymmetry factor into sensitivity to image change? Our data does not allow us to conclusively answer these questions. However, we speculate that an interesting possibility is that object motion is used as a "teaching signal" for refining an observers' ability to detect subtle image-level appearance changes (Balas & Sinha, in press), but that this process is "gated" by whether or not a motion sequence offers new information for prediction. That is, refining one's ability to say that image A differs from image B may be dependent on first learning which image is likely to come first in a sequence. This could explain why backward training leads to lower performance than forward training in Experiment 2. Forward walking could induce a time-symmetric sensitivity increase that "washes out" the pre-existing predictive effect. By contrast,

exposure to backward walking might lead to the "unlearning" of forward prediction (and/or learning of backward prediction) before sensitivity can increase.

We acknowledge that the above proposal is presently pure conjecture, and offer it only as an example of how symmetric and asymmetric learning may be intertwined in a conceptual model of dynamic object learning. The current data makes a strong case for an influence of temporal asymmetry on object perception over both long (Experiment 1) and short (Experiment 2) timescales, but does not provide enough data to constrain the mechanism by which such asymmetries contribute to the ultimate representation of object appearance. Our main contribution is to highlight the need for models of object learning to incorporate both symmetric and asymmetric spatiotemporal information. Furthermore, our results suggest many interesting avenues of further inquiry that may lead to a richer understanding of how prediction and association contribute to object learning in a dynamic setting. We conclude by briefly mentioning some of these possibilities for future work.

One important avenue for further investigation would be to examine the role of symmetric and asymmetric spatiotemporal input using a set of novel objects. This would offer the experimenter the opportunity to control the extent of temporal asymmetry during initial training, measure behaviour over extended posttraining intervals, and test the generality of any observed effects over multiple stimulus transformations. This would allow the time course (and consequences) of predictive and associative learning to be specified in much finer detail. A second important issue is the persistence of the training effect we have observed in Experiment 2. Over how long an interval does exposure to a moving object exert an effect on sensitivity? How do factors such as the length of the training period and observers' attention to the object (rather than the full visual field) modulate the influence of a training sequence over a span of hours, days, or weeks? Also, another important extension of this work would be to consider how the observation of object motion influences the subsequent perception of moving objects, rather than static images. Measuring the effect of various types of training on the perception of apparent motion or motion detection would help further integrate classical models of object recognition with ongoing attempts to characterize dynamic object representations.

Finally, there are multiple questions to be asked regarding the specific effects reported here. To what extent does the exposure effect generalize across various transformations of the stimulus? For example, does exposure to a walker at one view affect change detection at another? Similarly, the generality of the effect with regard to walker gender, identity, gait, or mood could be tested. These manipulations could help specify the "tuning" of body and locomotion perception.

## CONCLUSIONS

We have presented evidence that spatiotemporal asymmetries in object motion can influence static object perception. The ability to detect subtle appearance changes is modulated by observers' exposure to the direction of object appearance change over both short and long intervals of time. Furthermore, asymmetric object motion has a broad impact on immediate perception, such that objects at multiple views and in rarely observed orientations are subject to the same effects. These results highlight the importance of object motion on form processing and suggest interesting empirical questions concerning the respective roles of prediction and association during dynamic object learning and recognition.

## REFERENCES

Balas, B., & Sinha, P. (in press). Observing object motion induces increased generalization and sensitivity. *Perception*.

Beck, M. R., Angelone, B. L., & Levin, D. T. (2004). Knowledge about the probability of change affects change detection performance. *Journal of Experimental Psychology: Human Perception and Performance*, *30*(4), 778–791.

Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review*, *94*(2), 115–147.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*, 433–436.

Chuang, L., Vuong, Q. C., Thornton, I. M., & Bülthoff, H. H. (2006). Recognising novel deforming objects. *Visual Cognition*, *14*(1), 85–88.

Cox, D. D., Meier, P., Oertelt, N., & DiCarlo, J. J. (2005). "Breaking" position-invariant object recognition. *Nature Neuroscience*, *8*(9), 1145–1147.

Cutting, J. E. (1987). Perception and information. *Annual Review of Psychology*, *38*, 61–90.

Diamond, R., & Carey, S. (1986). Why faces are and are not special: An effect of expertise. *Journal of Experimental Psychology: General*, *115*, 107–117.

Dittrich, W. H. (1993). Action categories and the perception of biological motion. *Perception*, *22*, 15–22.

Foldiak, P. (1991). Learning invariance from transformation sequences. *Neural Computation*, *3*, 194–200.

Freyd, J. J. (1983). The mental representation of movement when static stimuli are viewed. *Perception and Psychophysics*, *33*(6), 575–581.

Freyd, J. J., & Finke, R. A. (1984). Representational momentum. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *10*(1), 126–132.

Gauthier, I., & Tarr, M. J. (1997). Becoming a "Greeble" expert: Exploring mechanisms for face recognition. *Vision Research*, *37*, 1673–1682.

Husk, J. S., Bennett, P. J., & Sekuler, A. B. (2007). Inverting houses and textures: Investigating the characteristics of learned inversion effects. *Vision Research*, *47*, 3350–3359.

Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Perception and Psychophysics*, *1*, 201–211.

Knappmeyer, B., Thornton, I. M., & Bülthoff, H. H. (2003). The use of facial motion and facial form during the processing of identity. *Vision Research*, *43*, 1921–36.

Kourtzi, K., & Kanwisher, N. (2000). Activation human MT/MST by static images with implied motion. *Journal of Cognitive Neuroscience*, *12*(1), 48–55.

Kourtzi, Z., & Shiffrar, M. (1999). Dynamic representations of human body movement. *Perception*, *28*, 49–62.

Kozlowski, L. T., & Cutting, J. E. (1977). Recognizing the sex of a walker from a dynamic point-light display. *Perception and Psychophysics*, *21*, 575–580.

Lander, K., & Bruce, V. (2000). Recognizing famous faces: Exploring the benefits of facial motion. *Ecological Psychology*, *12*, 259–272.

Marr, D., & Nishihara, H. K. (1978). Representation and recognition of the spatial organization of three-dimensional shapes. *Proceedings of the Royal Society of London, Series B*, *200*, 269–294.

Moses, Y., Adini, Y., & Ullman, S. (1994). Face recognition: The problem of compensating for illumination changes. In J.-O. Eklundh (Eds.), *Proceedings of the European conference on Computer Vision* (pp. 286–296). Stockholm: Springer.

Newell, F. N., Wallraven, C., & Huber, S. (2004). The role of characteristic motion in object categorization. *Journal of Vision*, *4*(2), 118–129.

Nijhawan, R. (1994). Motion extrapolation in catching. *Nature*, *370*, 256–257.

Pavlova, M., & Sokolov, A. (2003). Prior knowledge about display inversion in biological motion perception. *Perception*, *32*(8), 937–946.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*, 437–442.

Pollick, F. E., Lestou, V., Ryu, J., & Cho, S. (2002). Estimating the efficiency of recognizing gender and affect from biological motion. *Vision Research*, *42*(20), 2345–2355.

Reed, C. L., Stone, V. E., Bozova, S., & Tanaka, J. (2003). The body-inversion effect. *Psychological Science*, *14*(4), 302–308.

Rosenblum, L. D., Yakel, D. A., Baseer, N., Panchal, A., Nordase, B. C., & Niehus, R. P. (2002). Visual speech information for face recognition. *Perception and Psychophysics*, *64*(2), 220–229.

Shiffrar, M., & Freyd, J. (1990). Apparent motion of the human body. *Psychological Science*, *1*, 257–264.

Sinha, P., & Poggio, T. (1996). The role of learning in 3-D form perception. *Nature*, *384*, 460–463.

Stone, J. V. (1998). Object recognition using spatiotemporal signatures. *Vision Research*, *38*, 947–951.

Stone, J. V. (1999). Object recognition: View-specificity and motion-specificity. *Vision Research*, *39*, 4032–4044.

Stone, J. V., & Harper, N. (1999). Temporal constraints on visual learning: A computational model. *Perception*, *28*, 1089–1104.

Ullman, S., & Bart, E. (2004). Recognition invariance obtained by extended and invariant features. *Neural Networks*, *17*, 833–848.

Verfaille, K. (1993). Orientation-dependent priming effects in the perception of biological motion. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*, 649–670.

Verfaille, K., & Daems, A. (2002). Representing and anticipating human actions in vision. *Visual Cognition*, *9*, 217–232.

Vuong, Q. C., & Tarr, M. J. (2004). Rotation direction affects object recognition. *Vision Research*, *44*(14), 1717–1730.

Wallis, G. (1996). Using spatio-temporal correlations to learn invariant object recognition. *Neural Networks*, *9*(9), 1513–1519.

Wallis, G. (1998). Spatio-temporal influences at the neural level of object recognition. *Neural Networks*, *9*, 265–278.

Wallis, G., & Bülthoff, H. H. (2001). Effects of temporal association on recognition memory. *Proceedings of the National Academy of Sciences*, *98*(8), 4800–4804.

Yin, R. K. (1969). Looking at upside-down faces. *Journal of Experimental Psychology*, *81*, 141–145.