Control of Visual Attention in the Context of Behavior

Mary Hayhoe
Center for Perceptual Systems
University of Texas at Austin

What controls selection of information from visual image?

Evidence for an approach to selection in terms of modules, rewards, uncertainty, and prior beliefs.

To understand human behavior, experimenter typically manipulates attention via instructions, or manipulates stimuli to capture attention.

In natural behavior, selection and timing under subject’s control.

Eye movements: overt indicator of attentional allocation

What underlies the momentary decisions of where/when to saccade?

Immediate behavioral goals govern attentional selection.

- where to attend, when to attend, and what information to get. (Plate for knife placement, jar rim for lid target, lid for pick-up.)

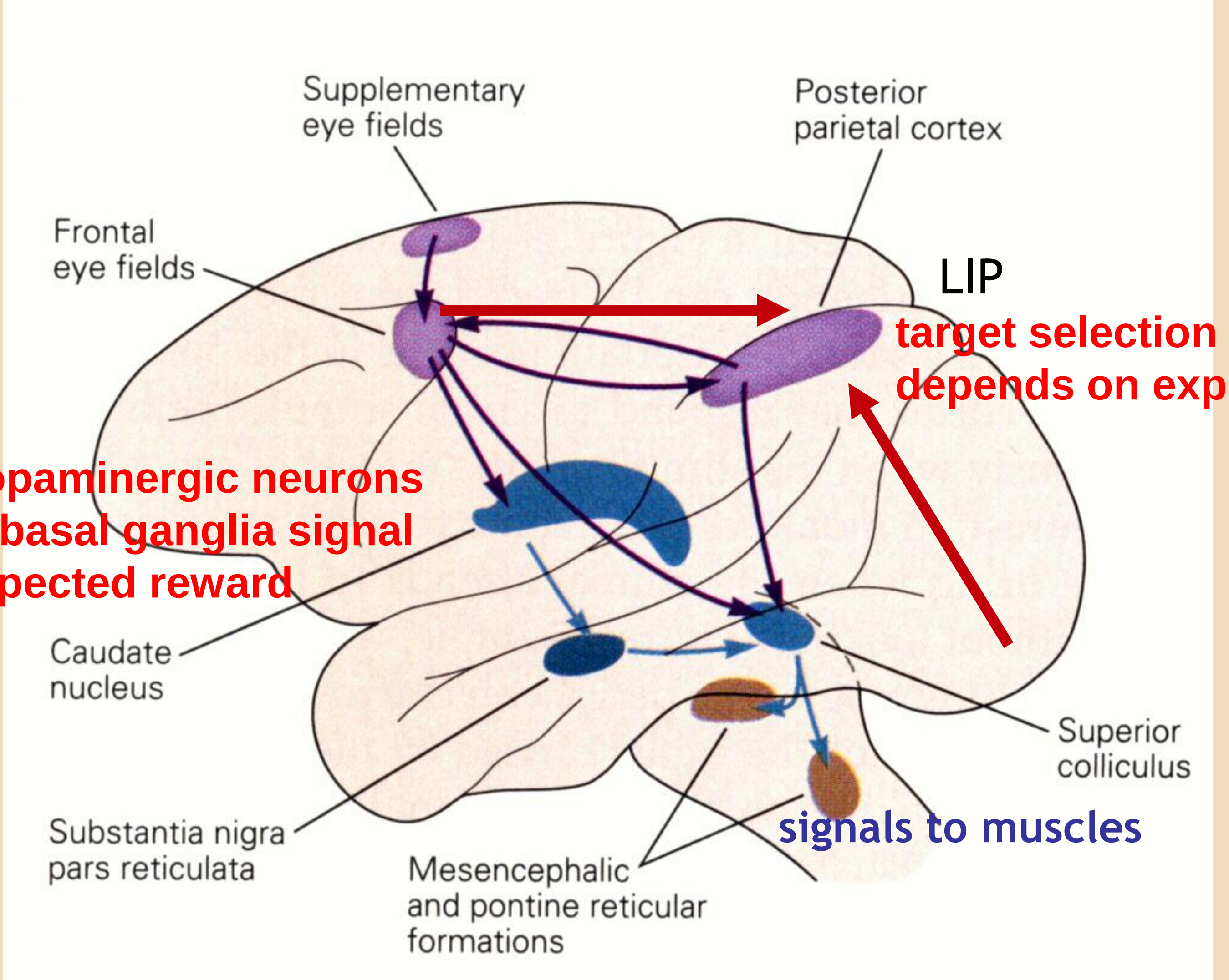
How does this come about?

Hypothesis: task control results from reward-based learning.

Reward sensitivity of Saccadic Circuitry

Neurons at all levels of saccadic eye movement circuitry are sensitive to reward.

Neural basis for reinforcement learning models of gaze behavior. (Schultz, 2000)



LIP

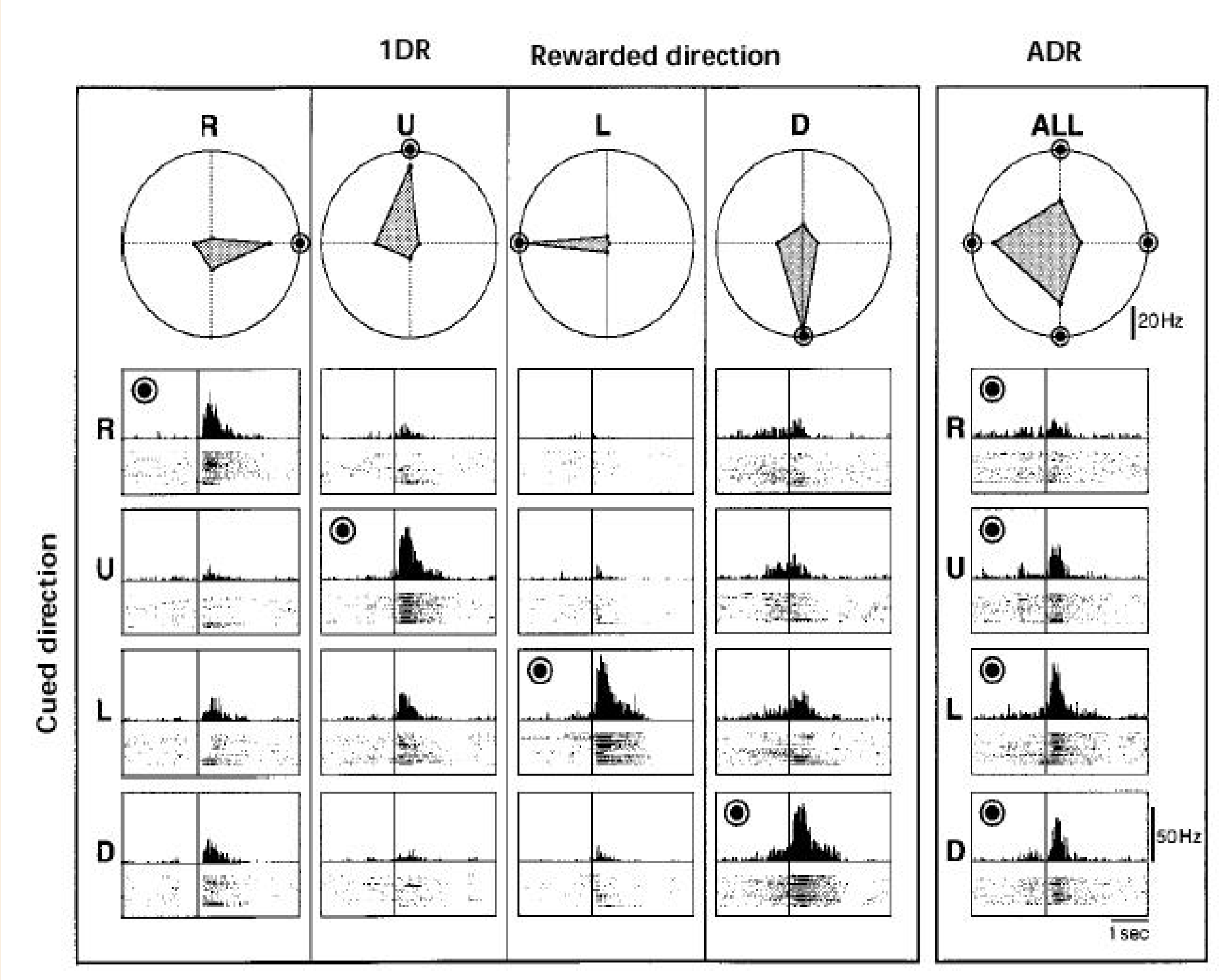
target selection depends on expected value

Dopaminergic neurons in basal ganglia signal expected reward

signals to muscles

Cells in caudate signal both saccade direction and expected reward.

Hikosaka et al, 2000



Monkey makes a saccade to a stimulus - some directions are rewarded.

R L Modeling of Gaze Control

Neural reward machinery provides a basis for RL models.

Walter the Virtual Humanoid

Sprague, Ballard, & Robinson TAP (2007)

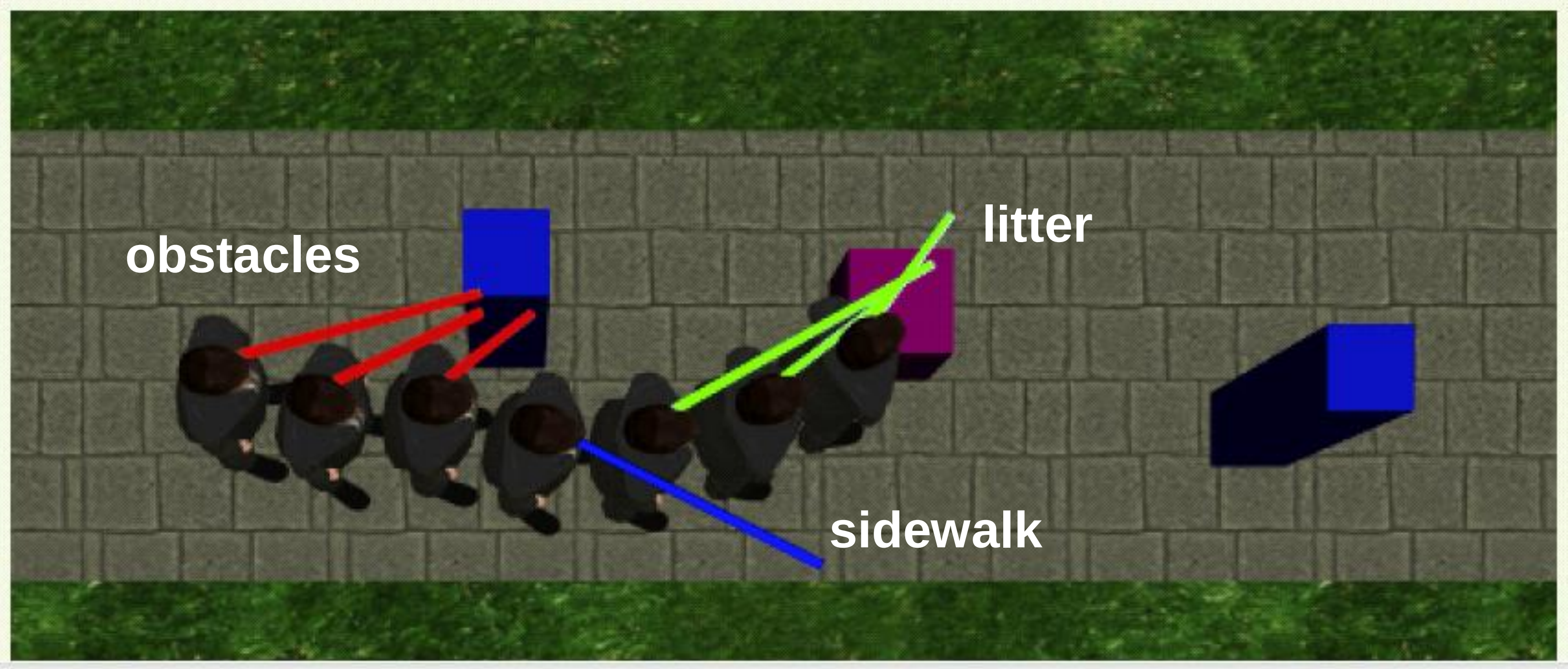
Virtual Humanoid has a small library of simple visual behaviors:

- Sidewalk Following
- Picking Up Blocks
- Avoiding Obstacles

Each behavior uses a ***limited, task-relevant*** selection of visual information from scene.

Behaviors have different priority/ reward value.

Controlling the Sequence of fixations



Gaze target is chosen based on both reward and uncertainty.

What about human behavior? Any evidence for role of reward?

Reward effects in neurons have been observed with very simple choice response paradigms eg “look to left target for high reward”.

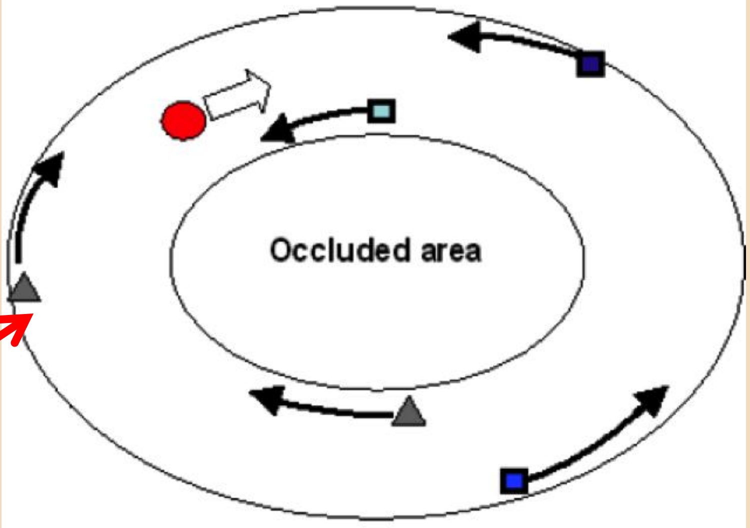
But eye movements are for getting information and are not directly rewarded in natural vision.

Need evidence for task (reward-based) control of gaze/attention in natural behavior.

Gaze allocation when walking in a real environment

Things to do: control direction, avoid obstacles, foot placement, characterize surroundings etc
Hypothesis: normal vision involves sets of sub-tasks or modules – need to allocate attention effectively between sub-tasks.

Portable ASL eyetracker
Oval path around large room



pedestrians

(Jovancevic & Hayhoe, 2009 J Neurosci)

How are gaze targets chosen?

Dynamic environments are tricky – timing of fixations more critical than in static scenes.

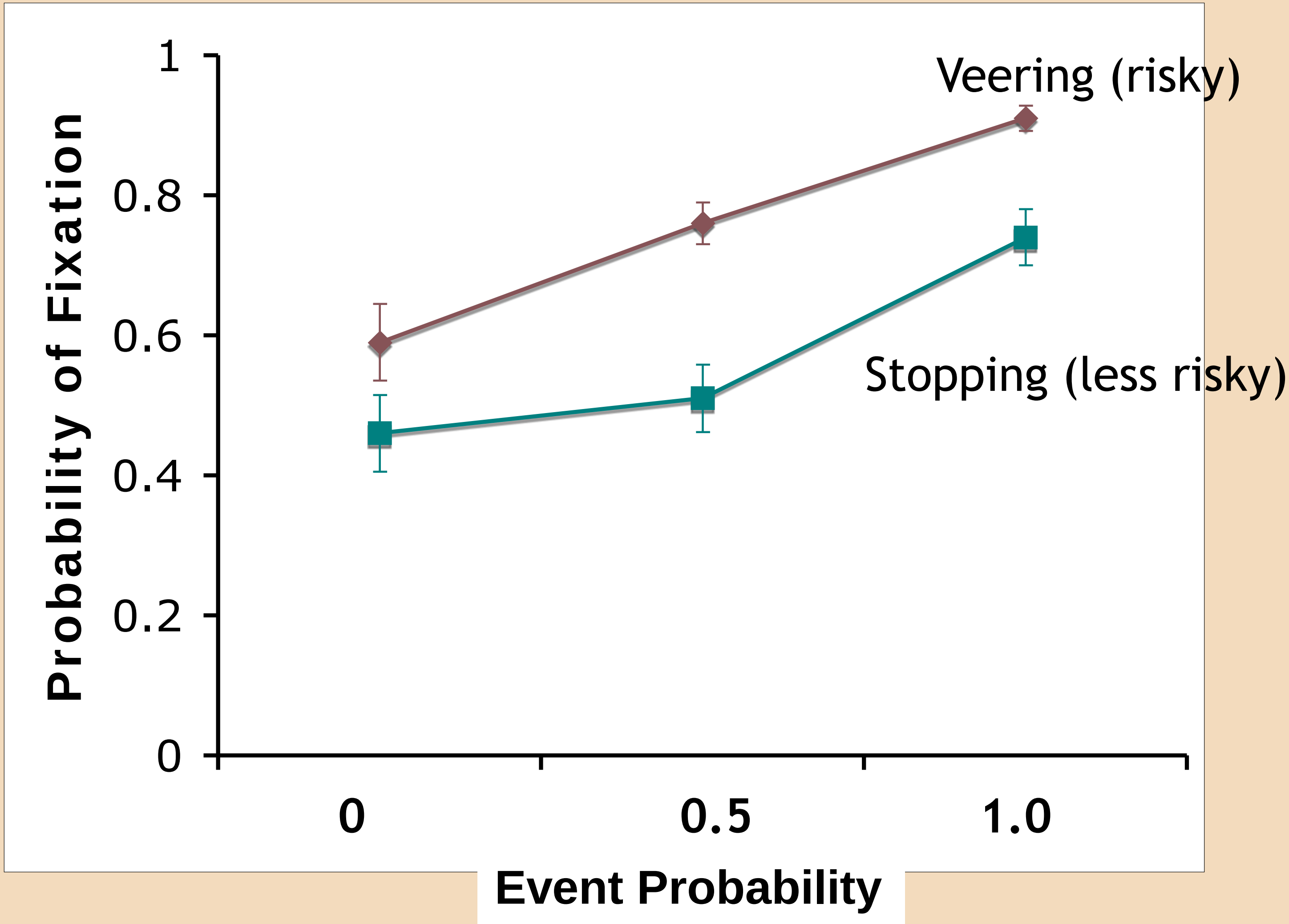
Manipulation of behavioral relevance (reward weight)

Occasionally some pedestrians either **stopped** for 1 sec or **veered** on a collision course with the subject

3 pedestrians behaved in characteristic ways:

Rogue pedestrian – always stops/veers
Safe pedestrian – never stops/ veers
Unpredictable pedestrian – stops/veers 50% of time

Fixation probability depends on behavioral relevance (subjective value) and probability of veering/stopping

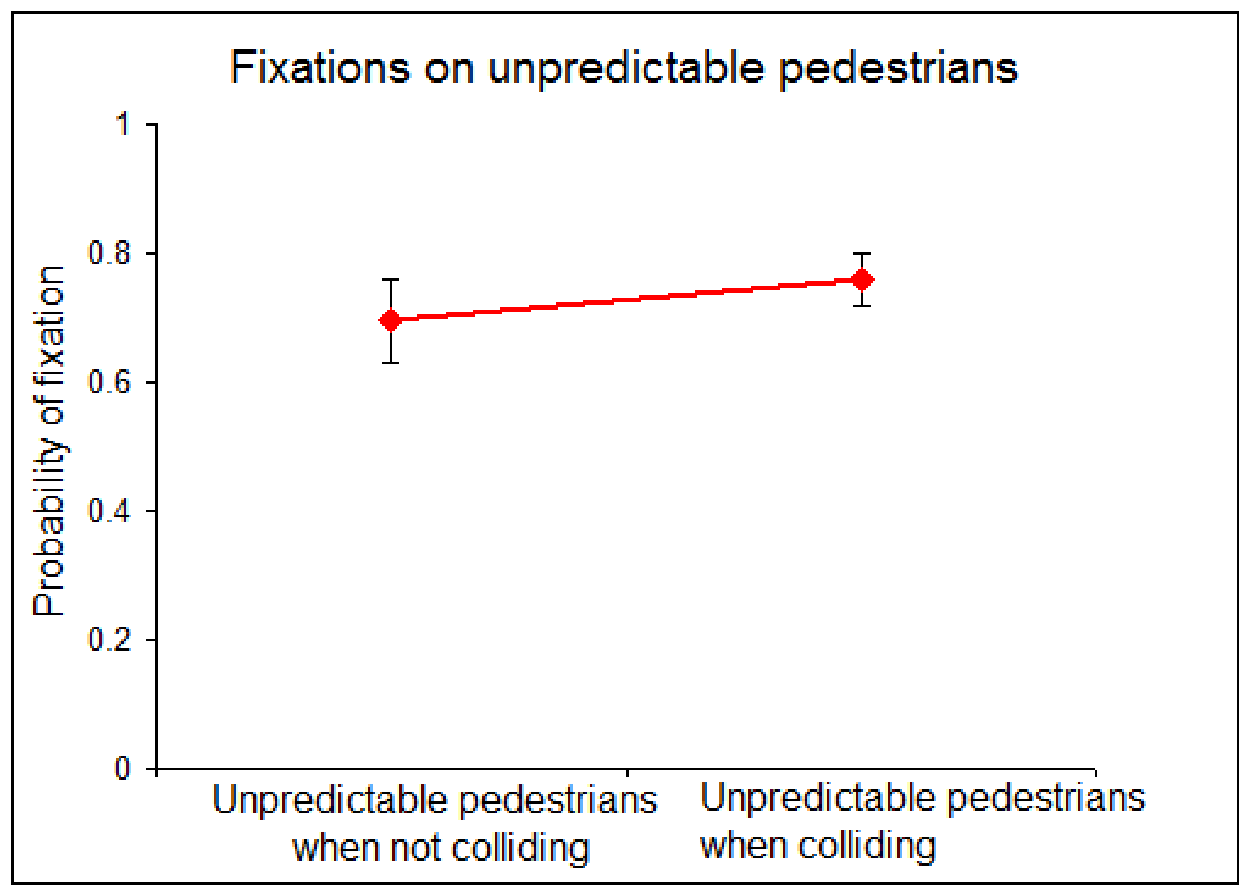


(Probability is computed during period in the field of view, not just collision interval.)

Almost all of the fixations on the Rogue were made **before** the veering onset (92%).

Thus gaze, and attention are anticipatory, based on history of events, not a result of attentional capture

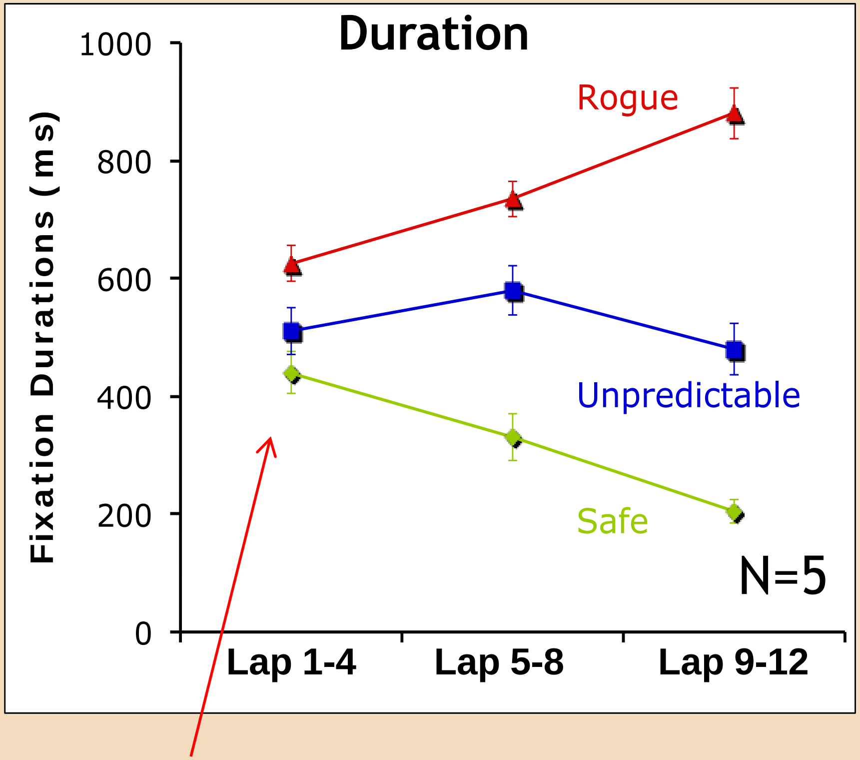
Gaze behavior based on expectation, not on veering event.



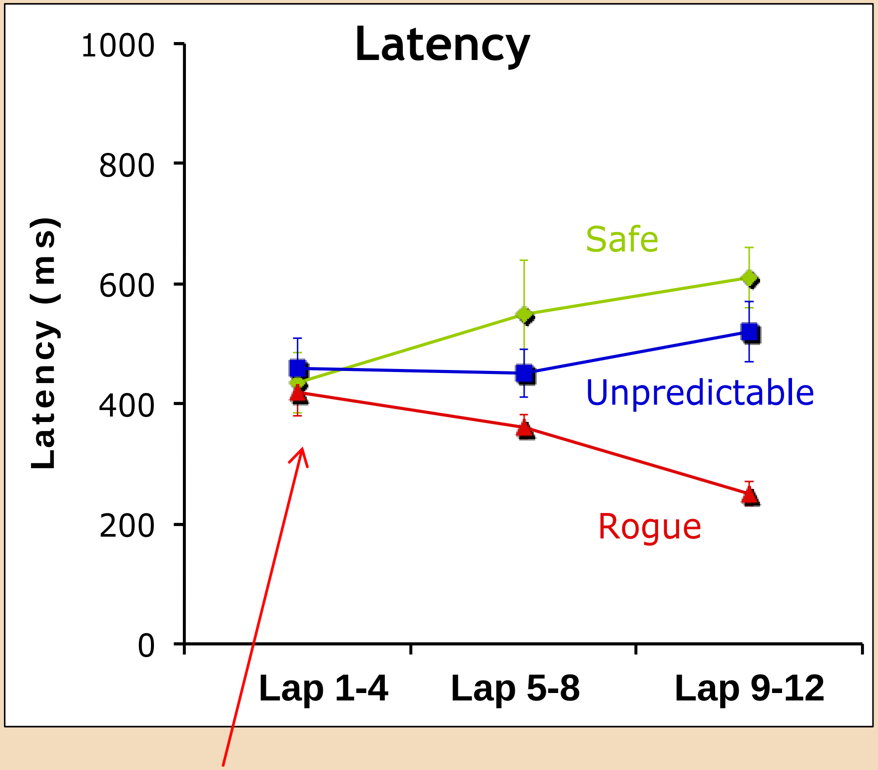
Probability of fixating unpredictable pedestrian similar, whether or not pedestrian actually veers on that trial.

Attention/gaze depends on reward probability (expected value)

Gaze behavior changes rapidly with experience (4-5 encounters): priorities re-allocated depending on behavioral relevance



prior



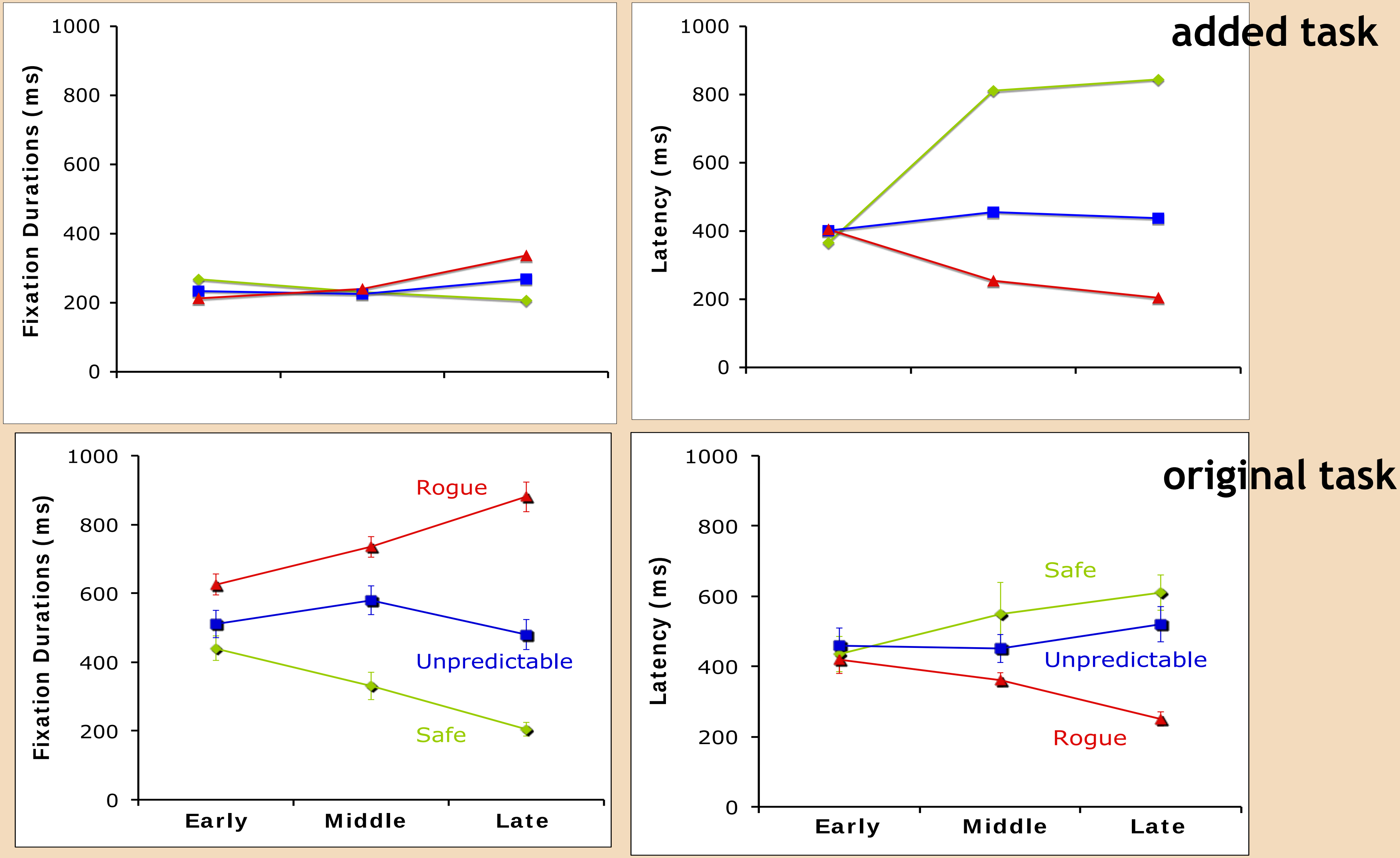
prior

Fixations on Rogue get longer/earlier, on Safe shorter/later

Sharing attention between tasks

Gaze priorities change when another task is added.

All fixations short duration, fixations on Safe deferred.



Fixations modulated by task importance/value (reward, and probability of reward). This implicates reinforcement learning models to account for control of attention and gaze.

Subjects learn the statistical structure of the world and allocate attention and gaze accordingly.

Control of gaze, and attention, is proactive, not reactive.

Subjects behave very similarly despite unconstrained environment and absence of instructions.

In walking paradigm, gaze behavior was anticipatory. Fixations on Rogue almost always occurred before veering event.

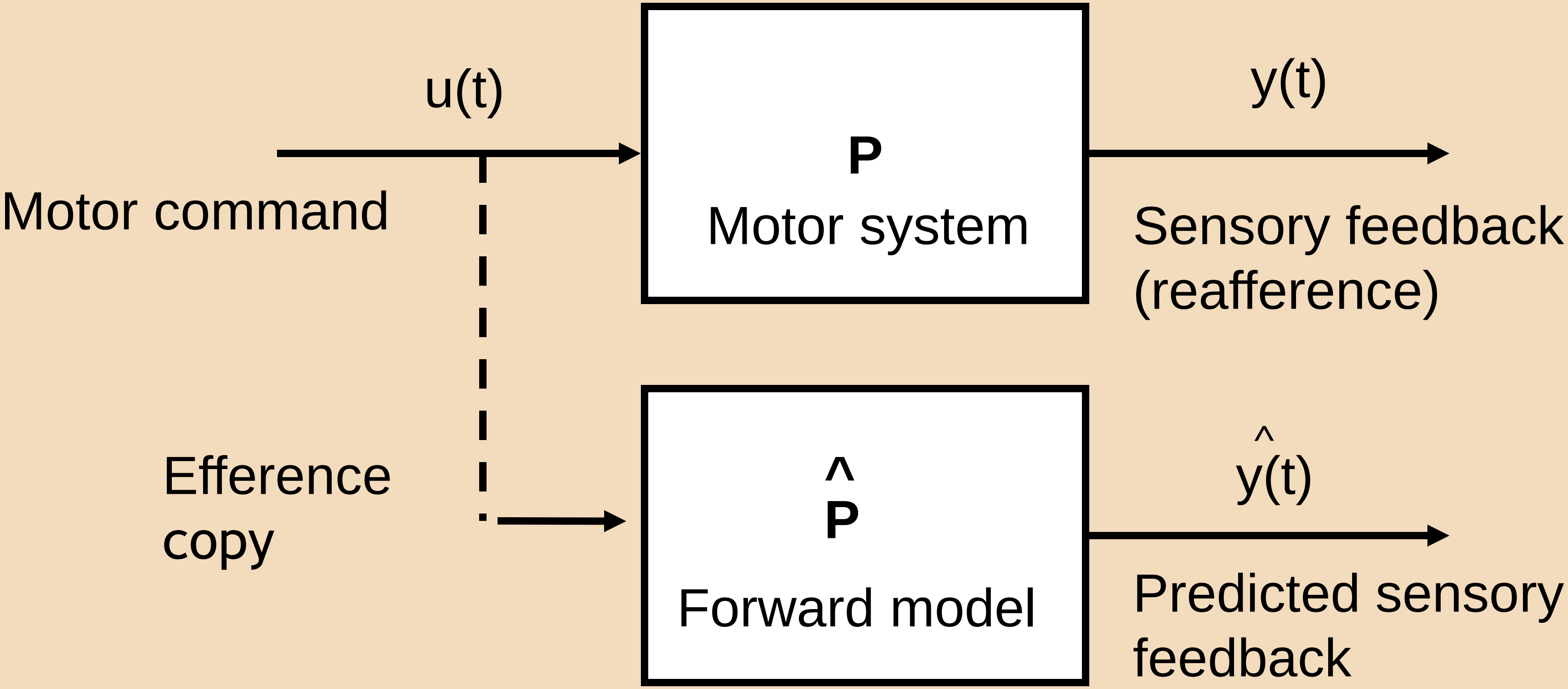
Fixations were based on predicted behavior.

Why?

Sensory delays make early planning of eye movements important/necessary.

What is the basis of prediction? Idea of Internal Models

In the case of body movements, forward models of body’s dynamics predict somatosensory consequences of movements (Wolpert et al, 1998).



Rapid comparison of actual with expected feedback circumvents delays

Internal models of visual world?

Evidence to the contrary..

Many actions can be controlled by the momentary visual signals in the image. (Warren, 2006)

For example, use looming information to compute “time-to-contact” to control interception/braking; “focus of expansion” to control heading.

That is, extract a “control variable”

Advantage: computational efficiency.

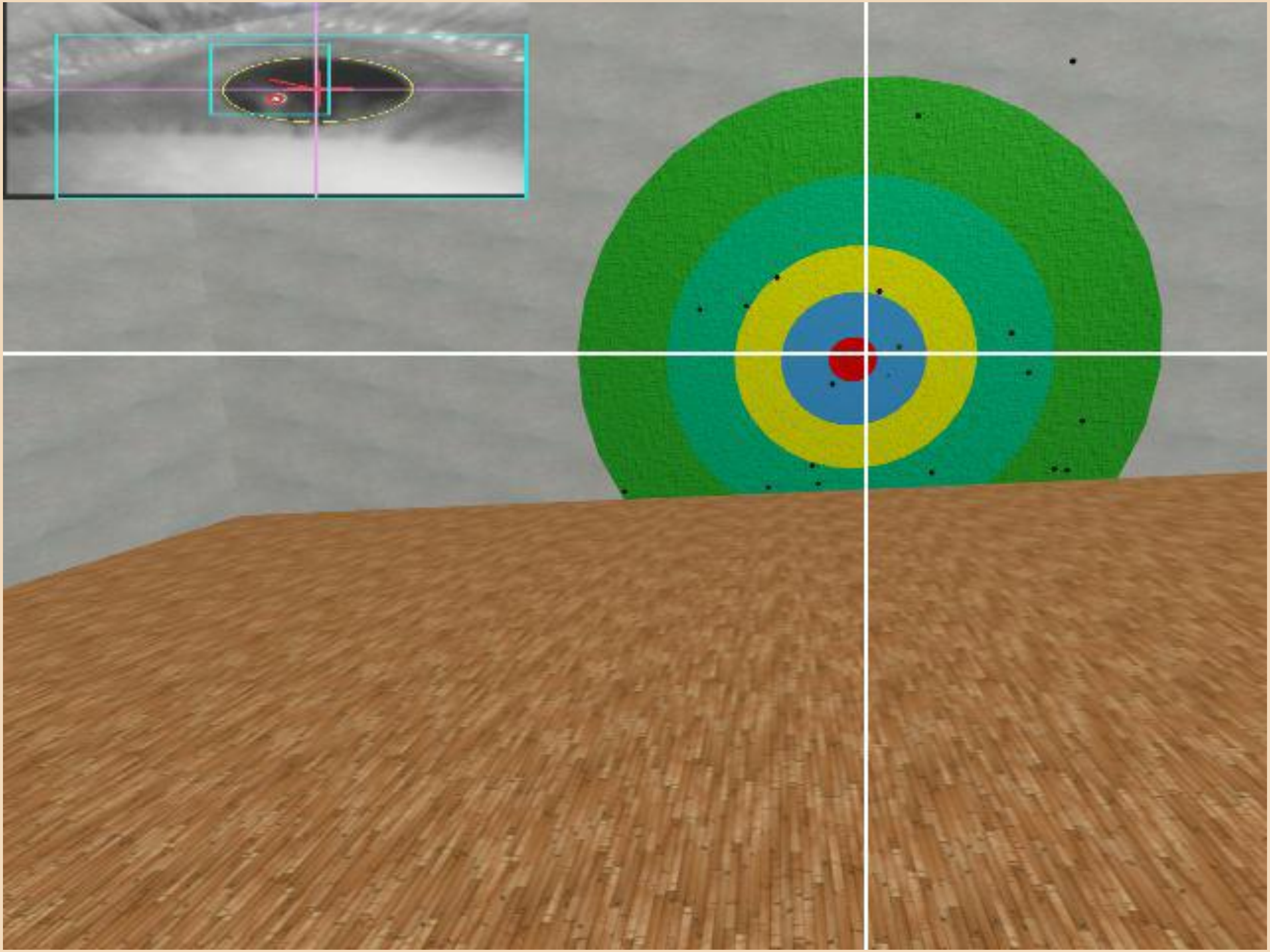
In natural movements, do we need internal models of visual image to generate eye movements in advance of events in the scene?

What is the nature of these internal models?

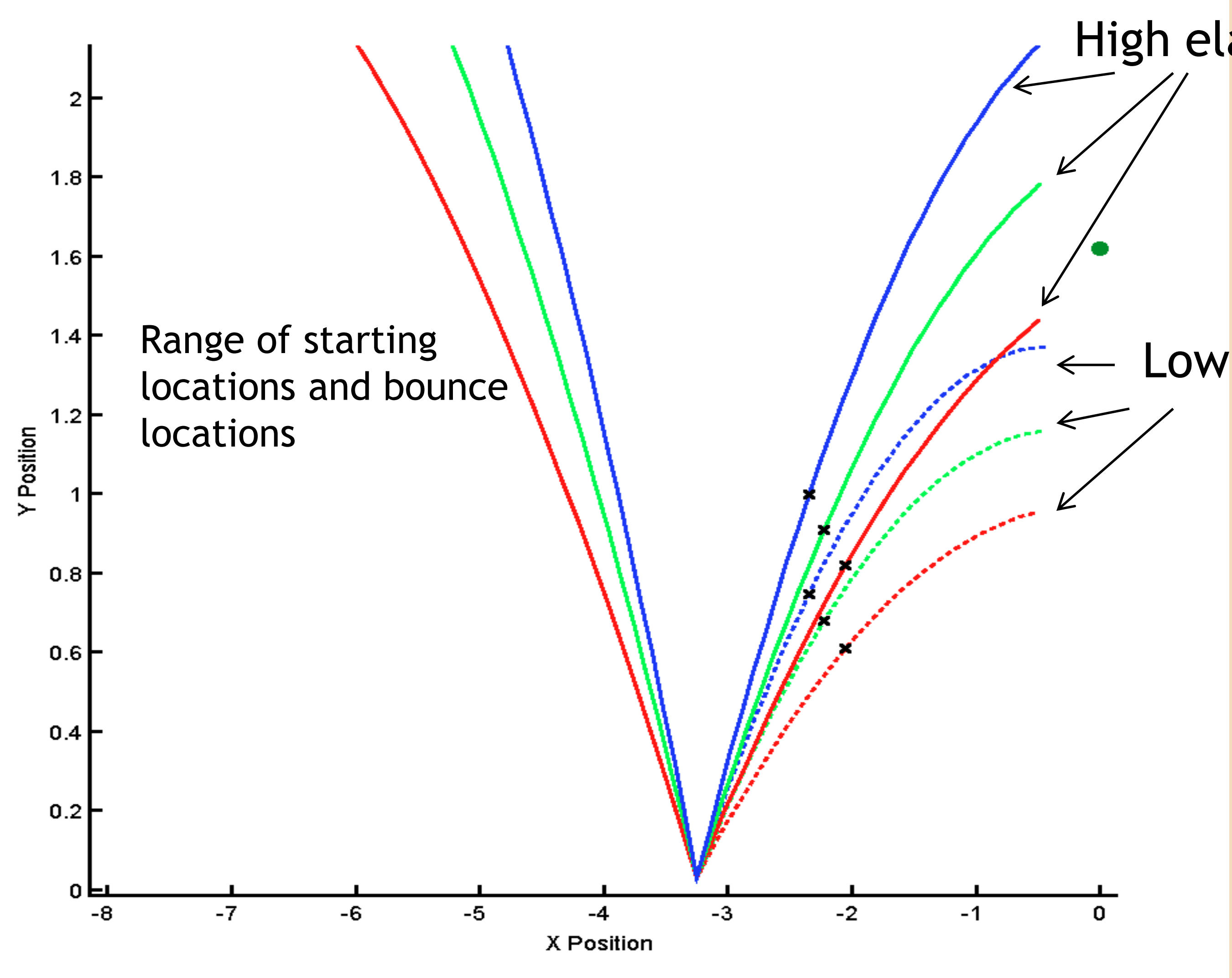
Gabe Diaz

Virtual racquetball:

Nvis helmet, Arrington eye-tracker, PhaseSpace head/hand/racquet tracking, Open Dynamics Engine to control ball and racquet interactions



Balls varied in vertical **velocity** and **elasticity**.
Velocity varied from trial to trial, elasticity was constant within a block.
Height after bounce predictable from current trial and previous trials within a block.

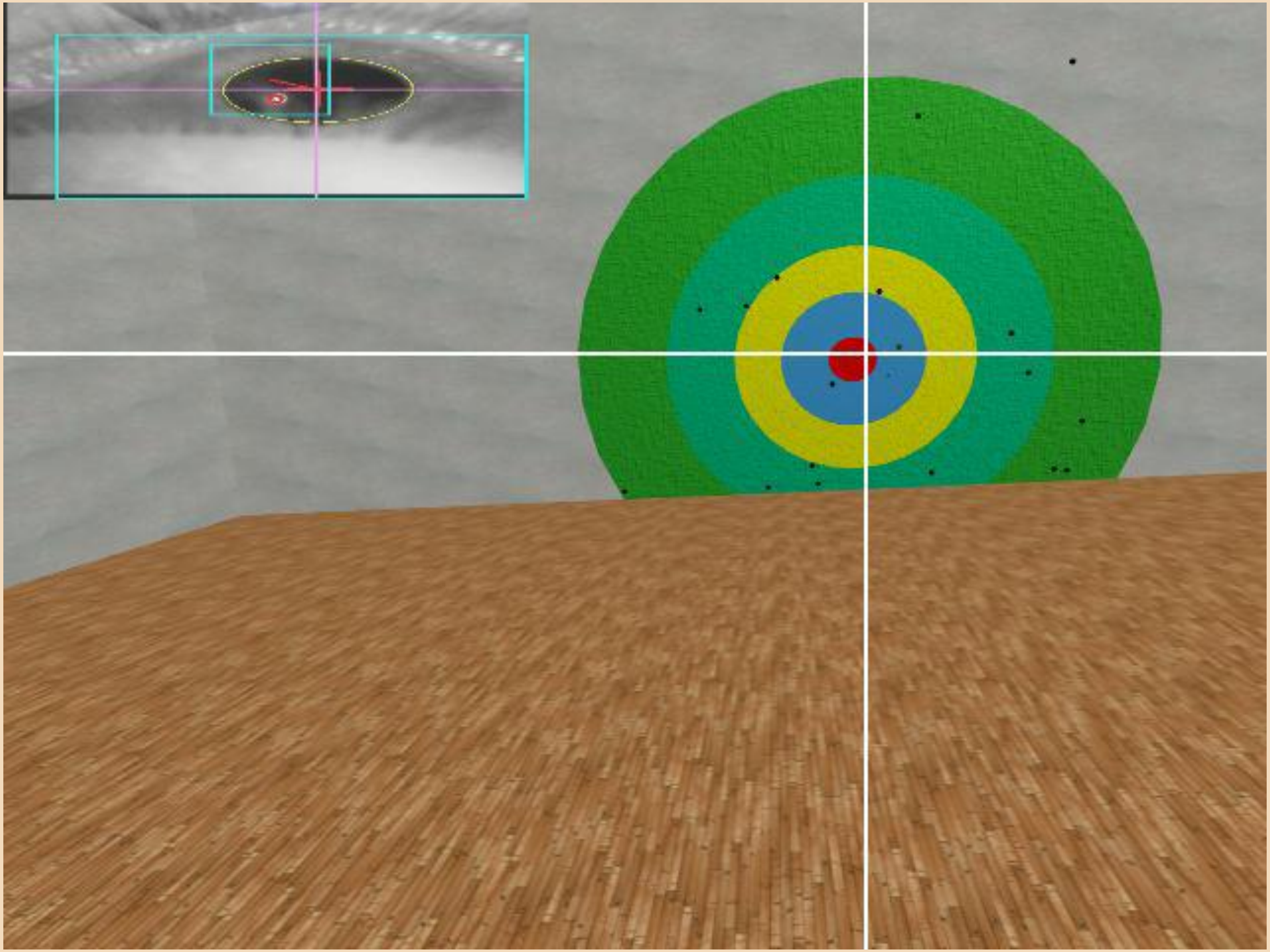


High elasticity

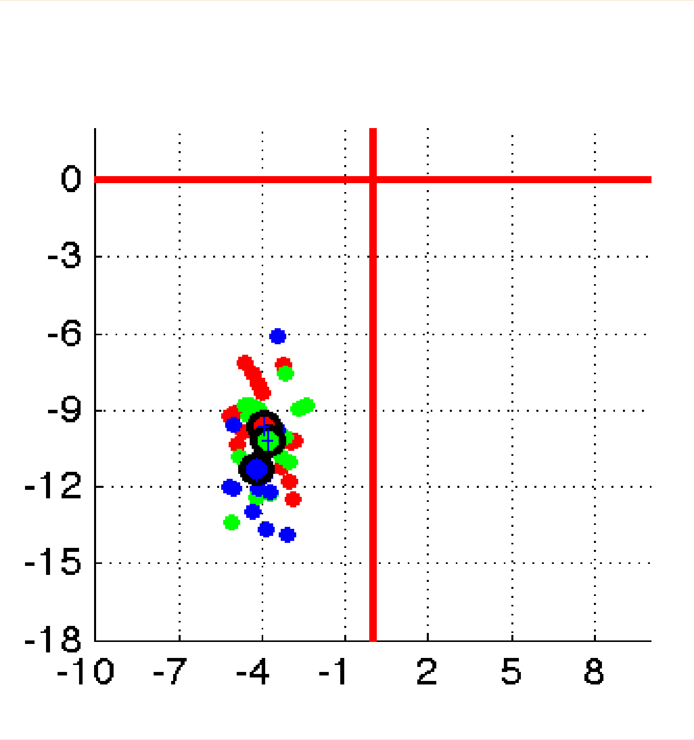
Low elasticity

Range of starting locations and bounce locations

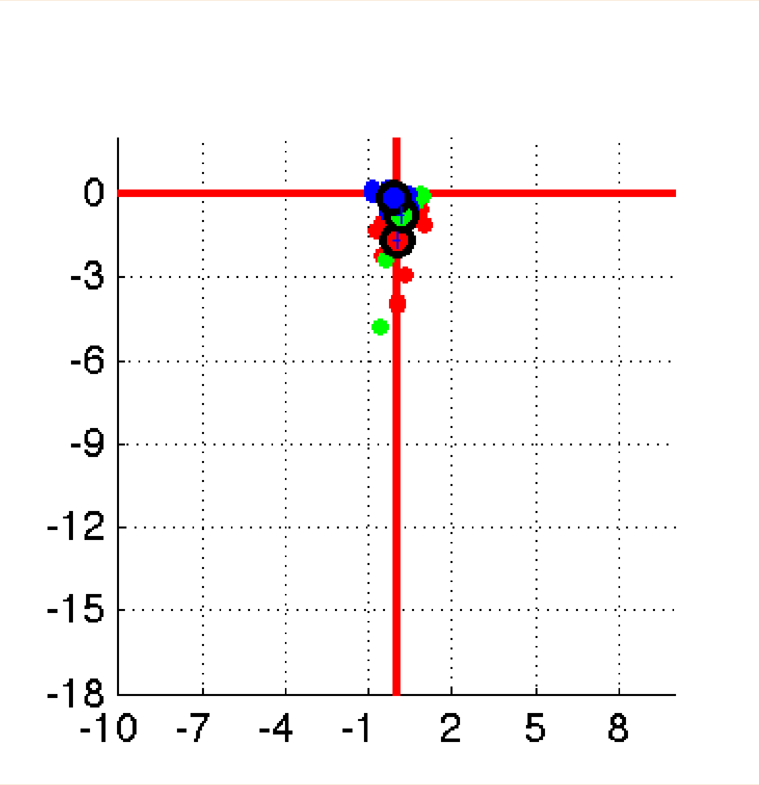
Saccade to a location ahead of the bounce



Subjects' gaze predicts location of ball after it bounces. Prediction is based on knowledge of elasticity, (based on history) plus velocity.



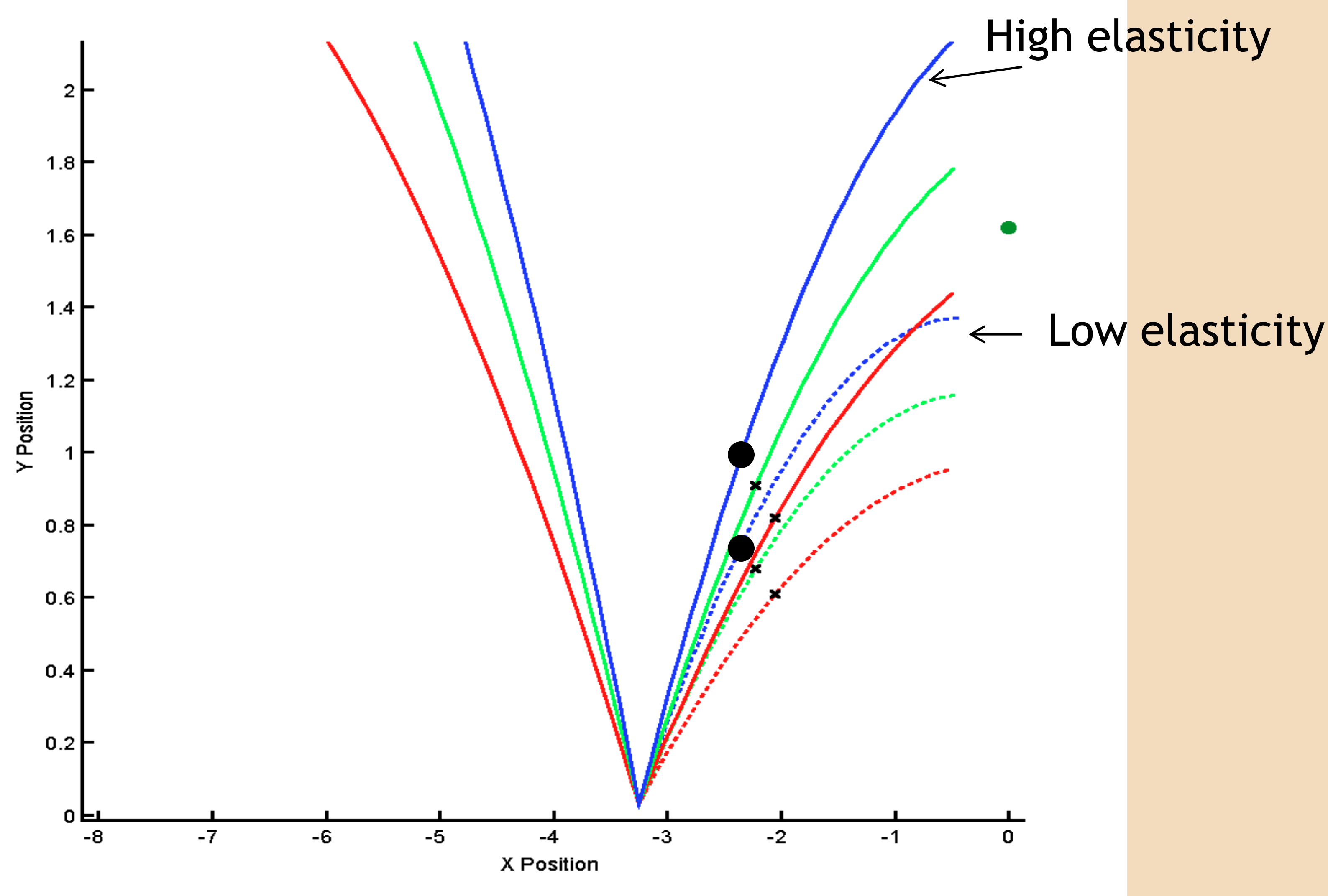
Ball location relative to gaze at time of bounce



Ball location relative to gaze 150 msec later.

Ss adjust predicted gaze point for elasticity and velocity.

Location of saccade precisely adjusted to compensate for elasticity and pre-bounce velocity

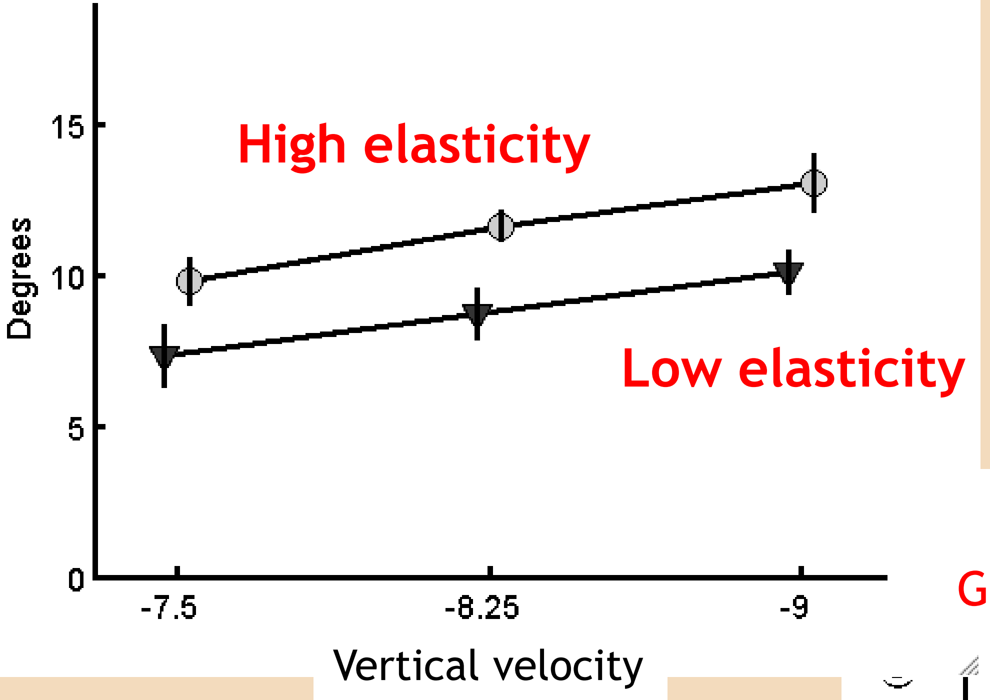


High elasticity

Low elasticity

Predictive Saccades: Location

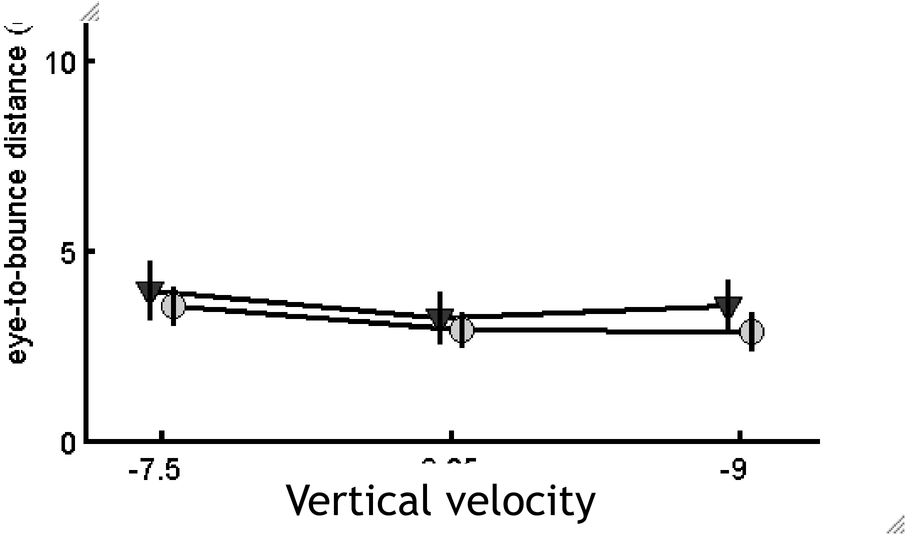
Gaze to ball distance at time of bounce



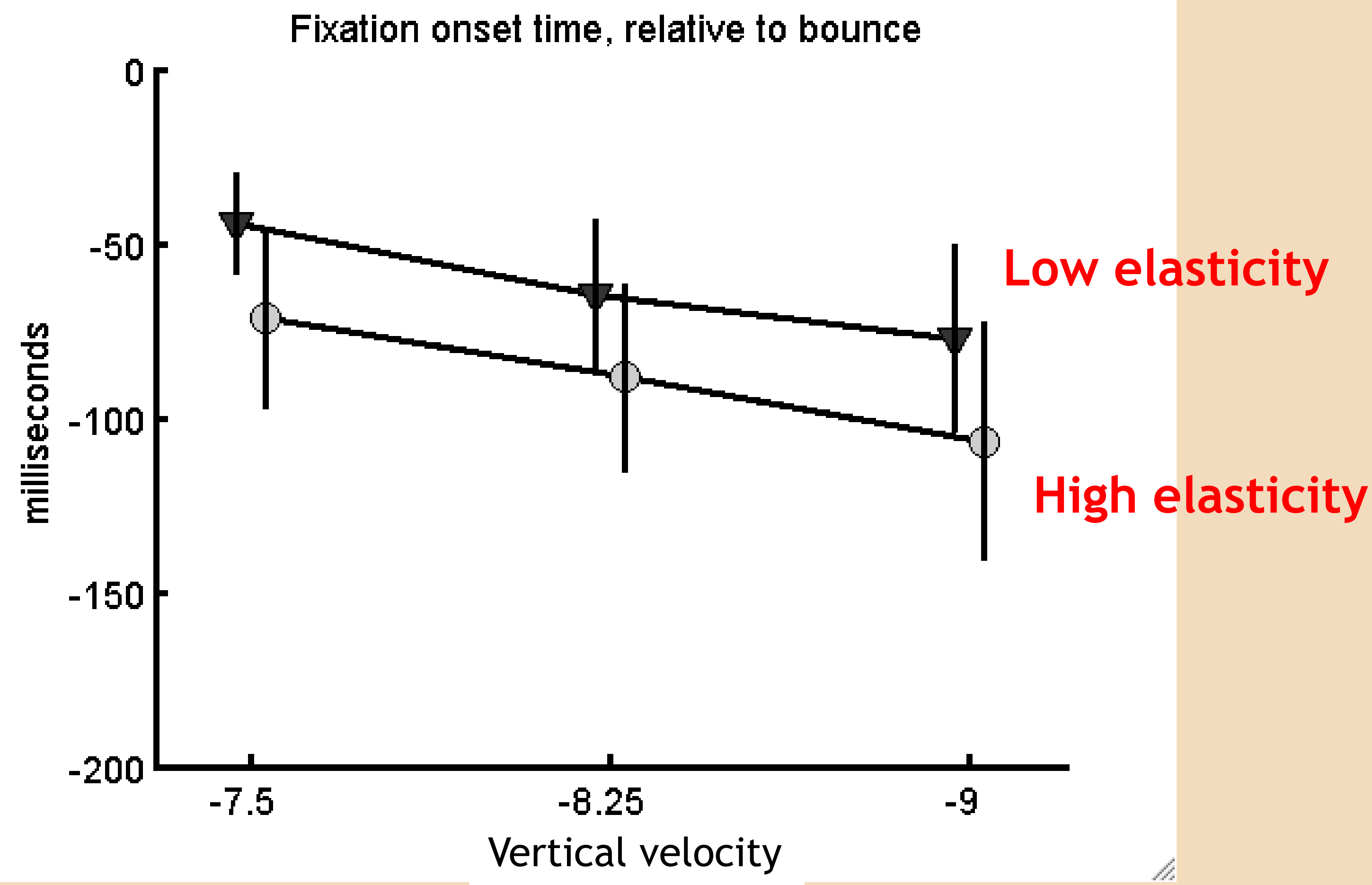
Subjects saccade to location above the bounce point.

Ball then passes close to gaze. Location of saccade precisely adjusted to compensate for elasticity and pre-bounce velocity

Gaze to ball distance at minimum

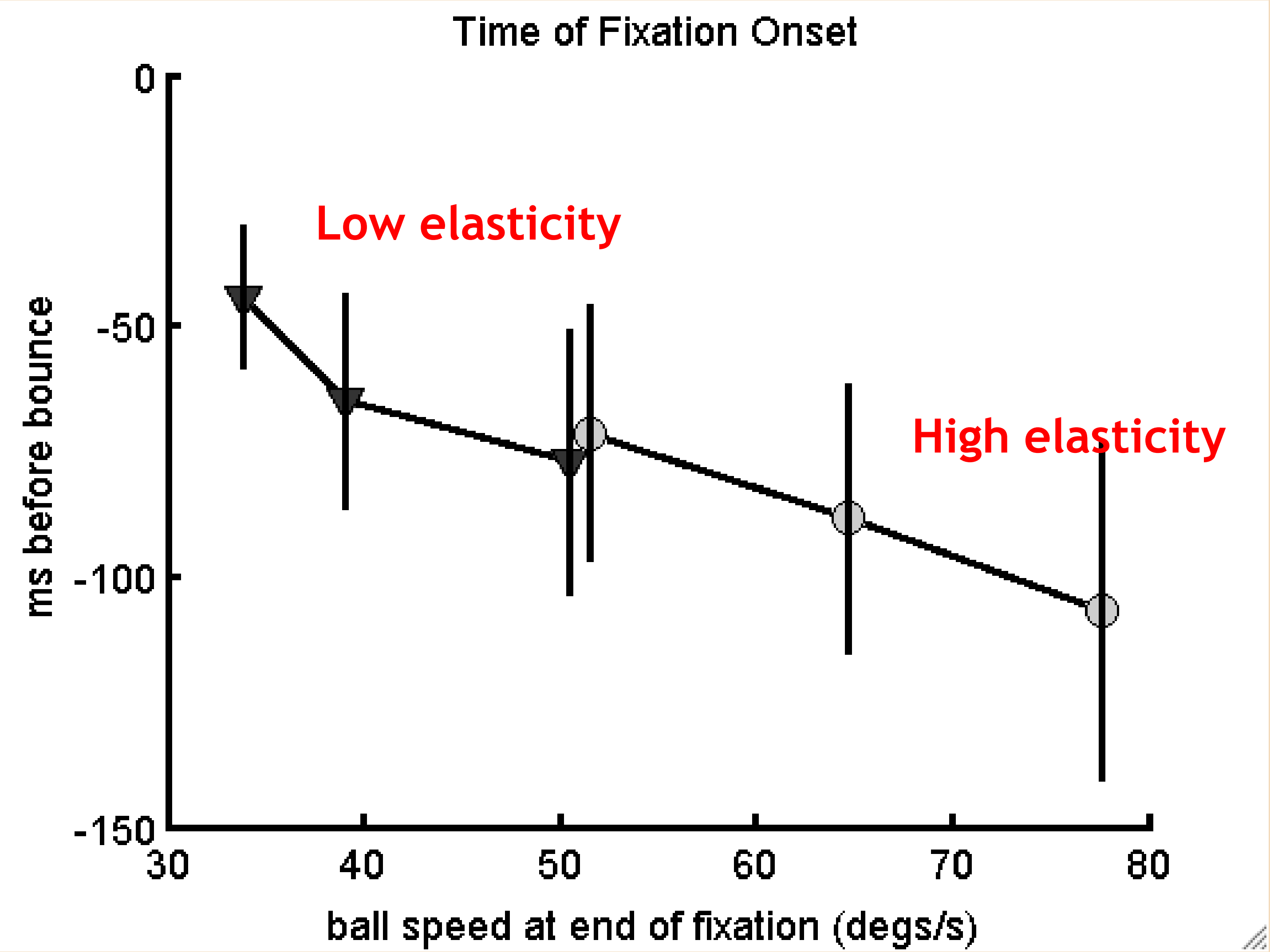


Predictive Saccades: Timing



Earlier saccade for more elastic balls (prior trials).
Earlier saccade for high velocity balls (current trial).

Predictive Saccades: Timing



Time of saccade related to post-bounce speed

Internal Visual Models Allow Prediction

Anticipatory saccades reveal that gaze is planned for a predicted state of the world.

Predictions must be based on some kind of internal model of visual events.

Note that predictions are very precise and similar between subjects, so model is very good!!

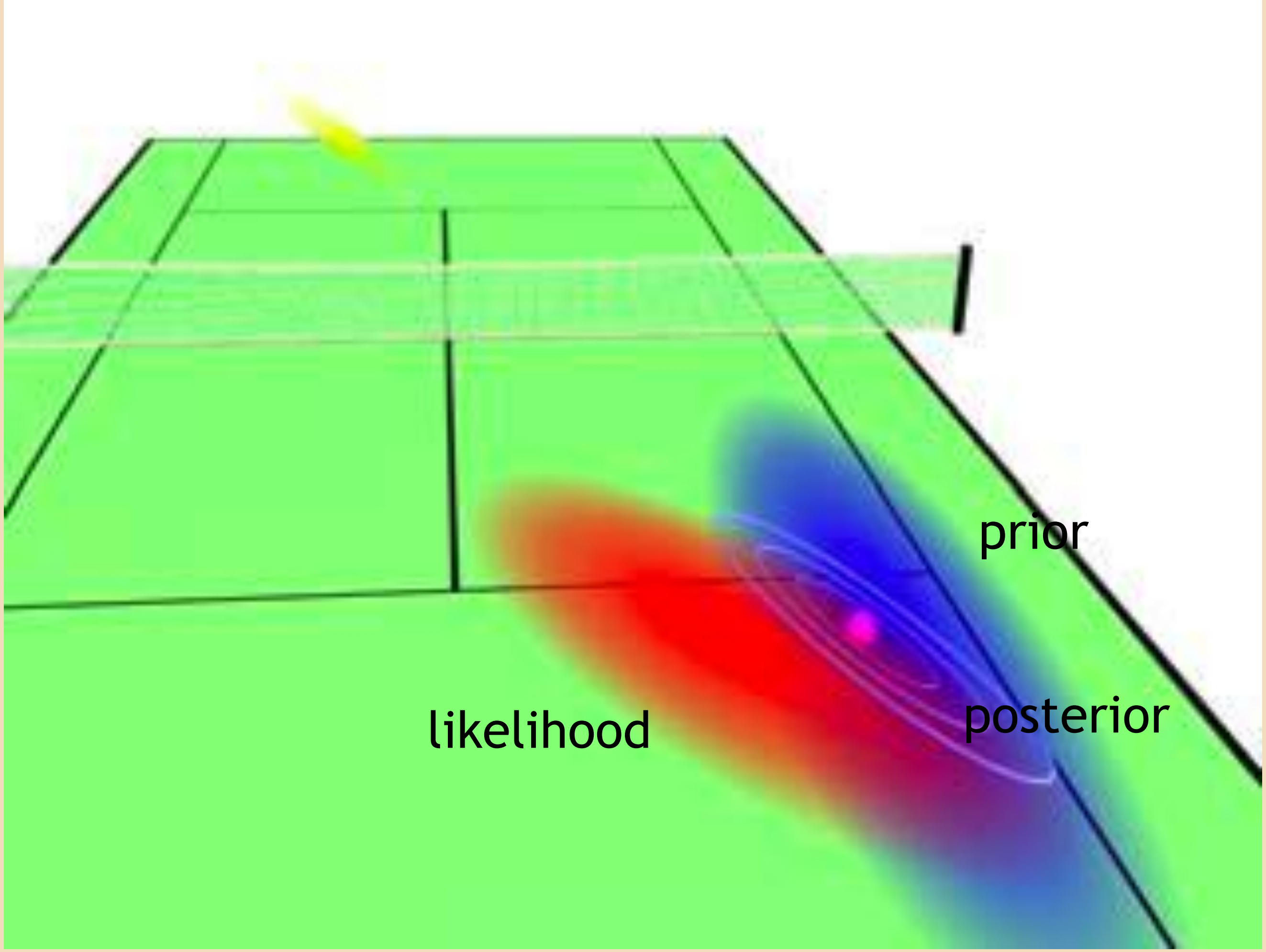
Internal Models

What do we know about the internal model? Evidence suggests it is high level and complex (angle, speed, elasticity, 3D, gravity, viewpoint independent).

In addition to mitigating the problem of visual delays, another value of experience-based internal visual models is that it allows better coordination of eye, head, arm, and body movements.

In reaching, evidence for the optimal Bayesian integration of current visual information with visual priors, (Koerding & Wolpert, 2004; Brouwer & Knill, 2007; Tassinari et al, 2006)
Perhaps a similar optimal weighting occurs with saccadic eye movements.

Hypothesis: Bayesian prediction of future state



Wolpert, 2005

Summary

Complex behavior can be broken down into sub-tasks or modules. This is consistent with observations of natural behavior.

Execution of sub-tasks/modules is learned and is governed by reward. Supported by gaze allocation in walking.

Learnt statistics/ priors about world state govern allocation of attention. Supported by both walking and racquetball.

Acknowledgments:

Jelena Jovancevic and Brian Sullivan – walking

Gabe Diaz – virtual racquetball