Using Sequential Structure to Improve Visuomotor Control

Anna Ma-Wyatt (anna.mawyatt@adelaide.edu.au)
School of Psychology, University of Adelaide, SA 5005, Australia

Daniel J. Navarro (daniel.navarro@adelaide.edu.au)
School of Psychology, University of Adelaide, SA 5005, Australia

Abstract

In daily life, people make rapid, goal-directed movements to interact with their environment. Since these movements are goal-directed, the outcome of the movement is important. A plan is typically formulated to make the movement using visual information about target location before the movement is initiated. However, in dynamic environments people need to track the location of an object if it moves during the reach. Additionally, it would be beneficial to motor performance to learn the distribution of target locations over multiple reaches. In this paper we develop a simple model that describes how people might exploit the sequential structure over a series of trials to improve rapid visuomotor control. We then present empirical data from a sequential tracking task that investigates how people's knowledge of the location of objects is updated over trials to improve pointing performance. The model is able to predict when people will hit and miss targets with reasonable success. Interestingly, the results suggest that not only are people able to use sequential information but also look for it even when it does not exist.

Keywords: visuo-motor control; Bayesian models; sequential learning

Introduction

People regularly make goal-directed reaches to interact with the world, where the goal or target of these movements is usually visually defined. During the course of such a reach, visual information about the target location relative to the hand will necessarily change. In part, this is because the hand is moving, but it can also be because the target itself is moving. In some cases the target is essentially stationary (e.g., pressing the car stereo button) and thus easily predictable. In others, the motor plan must be adapted during the course of the movement, because the target can move in an unpredictable fashion (e.g., swatting a mosquito). To be able to reach accurately and precisely, people must therefore be able to code not only the location of the goal, but also be able to respond to any changes in information about the goal that might affect the outcome of the reach.

It makes intuitive sense that people develop beliefs about the probability of events occurring, and that these beliefs might be updated as the events actually occur. For example, rapid motor control in everyday life appears to show a strong effect of prior expectation. Consider the visuomotor control problems facing the "sandwich artists" at a delicatessen during the busiest time of day: most people who order sandwiches will ask for lettuce as one of the ingredients. As a consequence, it is not uncommon for the server to reach for the lettuce before it is requested. If one does not in fact subsequently order lettuce, this occasionally produces errors.

Taken together, these two observations (that targets can move, and that people's prior beliefs about targets matter) suggest that when people are asked to point to targets that can move in a potentially predictable manner, people will try to anticipate where the target is likely to appear, and use this knowledge to improve the accuracy of their pointing. In this paper, we explore this topic. We report an experiment and a simple model developed to investigate the extent to which people exploit the statistical background knowledge to alter movements under time pressure.

Influences on Rapid Pointing

There are several key factors that significantly influence rapid pointing performance. The quality of visual information about target location available during the planning stage can limit pointing precision (Ma-Wyatt & McKee, 2006); similarly, the expected error associated with the reach influences the planning, which in turn relies on visual information about where the hand landed relative to the target (Trommershäuser et al, 2005). Finally, the well known speed-accuracy trade-off for controlling movements (Fitts, 1954) means that time pressure can play an important role: without time pressure, people have ample time to integrate feedback and refine movements, and so will need to rely less on their prior beliefs.

More closely related to the current work is the finding that people make use of information differently depending on when it becomes available. In particular, the weight given to visual information is increased in the final stages of the movement, most likely because it tends to be the most informative about the outcome of the movement (e.g. Sober & Sabes, 2005). Moreover, there is some evidence that people can optimally integrate multiple sources of sensory input during a reach. For instance, Körding and Wolpert (2004) demonstrated that if there was increased uncertainty associated with new visual feedback during a reach, observers relied more heavily on prior beliefs about sensory estimates, integrating these estimates in a manner consistent with Bayesian statistics. In short, it may be that the visuomotor system can adapt to changing sources of information over short time scales to achieve and maintain a consistent level of motor performance.

An interesting extension to this line of work is to consider how the information available to the observer changes over a time scale of multiple trials. For example, a recent experiment by Ma-Wyatt & McKee (2007) investigated people's ability to could correct their reach in an online fashion. If the

target location changed early in the reach, people were able to hit the new target accurately; if the target shifted late in the reach, they could not.[1] However, when the shift occurred late, people showed a sensitivity to prior probabilities. When the target was equally likely to appear at the two possible locations, observers pointed to a location approximately equidistant from each. When the relative probability of the target landing at one location was altered, then the mean pointing location was shifted towards the more likely location.

Prior Beliefs & Rapid Movements

In view of this literature, it seems likely that people's solutions to rapid visuomotor inference problems are strongly influenced by the prior beliefs they hold about the relative likelihoods of different outcomes. With this in mind, consider the problem faced by a learner who – as part of their participation in an experiment – encounters a series of observations $\boldsymbol{x}_n = (x_1, \ldots, x_n)$. What prior beliefs might he or she bring to the next trial in the experiment, x_{n+1}? Should the learner average across all previous trials in the experiment, or weight some observations more heavily than others? If the latter is true, it might be expected that more recent past events are more predictive of the current experiences. If this occurs, how far back in time do people integrate past observations in order to make predictions, and does this change as a function of the statistics of the task itself? In the spirit of recent cognitive theories of sequential prediction (e.g., Brown & Steyvers 2009; Yu & Cohen 2009) we describe a simple Bayesian model that is applicable to rapid pointing tasks.

Bayesian Beliefs & Non-Stationary Worlds

Suppose that the object to be tracked may occur at one of m distinct locations, $z_1 \ldots z_m$. Letting $\theta_k = \Pr(x = z_k)$ denote the observer's subjective belief that an object x will appear at the kth location, we might describe the belief state by $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_m)$, and specify a symmetric Dirichlet(α) prior over possible values for $\boldsymbol{\theta}$. Formally, the observer should assume the following generative model for the first observation,

$$\begin{array}{rcl} x_1 & \sim & \text{Multinomial}(1, \boldsymbol{\theta}) \\ \boldsymbol{\theta} & \sim & \text{Dirichlet}(\alpha, \ldots, \alpha) \end{array} \tag{1}$$

If the probabilities remain invariant across trials, then standard Bayesian learning implies that, having observed the data $\boldsymbol{x}$, the posterior distribution over $\boldsymbol{\theta}$ would become

$$\boldsymbol{\theta} \,|\, \boldsymbol{x} \sim \text{Dirichlet}(\alpha + n_1, \ldots, \alpha + n_m) \tag{2}$$

where n_k counts the number of previous occasions on which a target fell in location z_k. After integrating out his or her uncertainty about the exact probabilities $\boldsymbol{\theta}$, the predicted probability that the next observation falls in location z_k is simply

$$\Pr(x_{n+1} = z_k \,|\, \boldsymbol{x}_n) = \frac{n_k + \alpha}{n + m\alpha} \tag{3}$$

(see Gelman, Carlin, Stern & Rubin, 1995).

In non-stationary environments, this updating method is not optimal (e.g., Arulampalam, Maskell & Gordon 2002; Yu & Cohen 2009). Suppose, for instance, that the predictive value of a previous observation decays exponentially over time (as is typical of simple "constant rate of change" processes). Then the belief updating method described previously will overweight old data relative to new data. This suggests that the knowledge available to the learner about $\boldsymbol{\theta}_{n+1}$, the probability distribution associated with the next observation, might be better modelled by

$$\boldsymbol{\theta} \,|\, \boldsymbol{x} \sim \text{Dirichlet}(\alpha + f(1), \ldots, \alpha + f(m)) \tag{4}$$

where $f(k)$ denotes a time-weighted sum of the relevant observations that fell in the kth location

$$f(k) = \sum_{i=1}^{n} \exp(-\lambda(n-i))\delta(x_i - z_k). \tag{5}$$

In this expression $\delta(\cdot)$ is the Kronecker delta function that is 1 if its argument is zero and 0 for all other input values. As a result the observer would predict that

$$\Pr(x_{n+1} = z_k \,|\, \boldsymbol{x}_n, \lambda) = \frac{\alpha + f(k)}{m\alpha + \sum_{i=1}^{m} f(i)} \tag{6}$$

For the present purposes, we fix $\alpha = .01$.

Movement Error Given Uncertain Beliefs

If an observer has expectations described by Equation 6, how might he or she construct motor plans for the pointing action, and what predictions does that imply about the probability that the observer will hit the target? To a first approximation, we may assume that Fitts' (1954) law holds, and the time t that an observer requires to make a reach to a target of width w located a distance d from the current one scales logarithmically,

$$t = a_1 + a_2 \log_2(2d/w). \tag{7}$$

If we assume that the learner starts with a plan to reach to location ℓ, and then partway through the reach discovers that the true target location is z_k, then Fitts' law provides a method to determine the form of the error probability as a function of distance. Making some simple assumptions about errors, Equation 7 suggests a simple model in which the implied hitting probability decays exponentially with distance:

$$\Pr(\text{hit} \,|\, \phi) = \exp(-\phi d(\ell, z_k)) \tag{8}$$

where $d(\ell, z_k)$ denotes the distance between locations ℓ and z_k. The free parameter ϕ in this expression is a function of the original parameters of Fitts' law (the width w is treated as fixed, while one of the two remaining degrees of freedom disappears because we assume that the observer can hit the target perfectly given an arbitrarily long time to do so).

If the observer knew the target location ahead of time, then the choice of motor plan is obvious: aim for the target. Given the uncertainty about potential locations, the story is a little more complex. One possibility is to choose a single plan that in some manner averages over this uncertainty; a second is to prepare multiple plans (one corresponding to each possible location), and to weight them accordingly. This loosely mimics the distinction between prototype models for categories

[1] Consistent with previous work, participants appeared to require approximately 150ms to update the movement plan online.

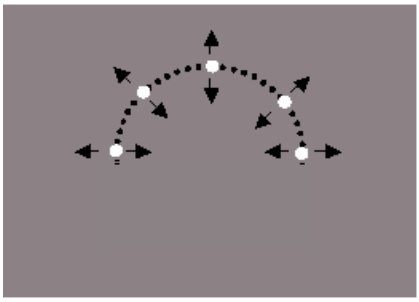

Figure 1: Schematic illustration of the display. Possible locations for the final target are shown as white dots, and the axis along which error was measured is shown by the black arrows. The intermediate target appeared at a point equidistant from all five locations (i.e., the centre of the circle), and the initial fixation point was below the intermediate one.

(e.g., Reed 1972) and exemplar models (e.g., Medin & Schaffer 1981), with the multiple plan model mapping onto exemplar models. Using an exemplar-style multiple plan model, the hit probability becomes

$$\begin{aligned}\Pr(\text{hit} \mid \phi, \lambda) &= \sum_{k=1}^{m} \exp(-\phi d(x_{n+1}, z_k)) \\ &\quad \times \Pr(x_{n+1} = z_k \mid \boldsymbol{x}_n, \lambda)\end{aligned} \tag{9}$$

The interpretation of the parameters in this model are as follows: ϕ governs the ability of the observer to compensate for errors during the course of the reach itself, while λ governs the extent to which the observer generalizes from old observations to new ones. Note that when $\lambda \to 0$, all observations are weighted equally and the belief updating reduces to a smoothed running average model (Equation 3). On the other hand, as $\lambda \to \infty$, all previous observations are assumed to be of zero predictive value, and so the observer has a uniform belief distribution on every trial. Only at intermediate values of λ does the observer show interesting sequential effects.

Experiment

We used a two step paradigm to investigate whether people could use sequential dependencies to update their position information. If it is the case that people are able to track sequential dependencies across trials, then we should observe an improvement in performance on consecutive trials, and a significant cost when the run of sequential dependencies ends. Our aim is to test whether people are able to use this information to improve pointing performance and to estimate over how many trials this information is retained.

Method

Participants. Four observers participated in the experiments. Three were naïve to the purposes of the experiment (MO, BS and MP); the other was an author (AM). A fifth observer (PH) also participated, but produced data too noisy to analyze. Additionally, due to time constraints, MP completed only half of the trials. All observers had normal or corrected to normal visual acuity and no known motor deficits. All observers gave informed consent to participate in the experiments.

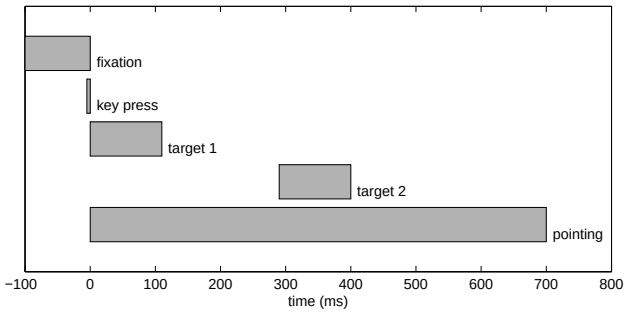

Figure 2: Time course of a single trial. The fixation dot remained on screen until a key is pressed, commencing the trial (0ms). The intermediate target appeared immediately, remaining on screen for 110ms before disappearing. The final location was revealed at the 290ms mark and vanished at 400ms. The observer was required to hit this target before the 700ms mark.

General procedure. The observer rested their chin on a chin rest and was seated 40cm away from a touchscreen in a darkened room. On each trial, the observer fixated a central point and made a key press on the keyboard in front of them to initiate the presentation of test stimulus. The target was a high contrast white dot that subtended 0.5° of visual angle (i.e., the dot was 0.7 cm in diameter). A white dot was presented for 110ms, after which it disappeared. After a delay of 180ms, another dot appeared in one of five possible locations, spread out in an arc as illustrated in Figure 1. The probability of the target appearing at any one of these locations was manipulated as outlined below. The target remained on the screen for 110ms (see Figure 2 for a timeline of an example trial). Observers were instructed to point as rapidly to the final location of the target as they could, and endpoints were registered by the touchscreen. We measured movement time as the time between the key press to initiate the trial, and the time at which the finger touched the screen.

Observers received negative feedback if their touch was too slow (>700ms), or too far away from the target location ($>2^\circ$ of visual angle, or ≈ 1.5cm). Because we were interested in determining if the observer would be able to use information about the sequential dependency of a target location across trials to alter their motor plan, it was important to ensure that reach times remained fast, and reasonably consistent across trials. We therefore used a points system to encourage observers to reach quickly. If the movement time was too slow, 150 points were deducted from the observer's score. If the point was within 2° of the target, the observer won 400 points. Points won were converted to a cash reward at the end of the experiment - observers were awarded 0.001¢ per point won.

Each observer completed several blocks of each of the sequential and non-sequential conditions. The order of sequential or non-sequential conditions was counterbalanced across observers. The order of the blocks for each of the sequential and non-sequential conditions was randomised for each observer.

Sequential structure. Each observer completed blocks of trials in which the position of the target was either sequentially dependent on the location of the target on previous trials (sequential) or randomly drawn (non-sequential). We cre-

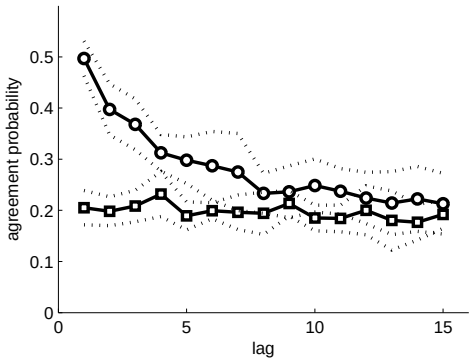

Figure 3: The sequential dependencies used in the experiment. Solid lines depict the average probability that a target location on trial t will fall in the same location as the target on trial $t - x$ for some lag x. For the autocorrelated data (circles), the agreement probability is 50% at lag 1 and then decays exponentially to a chance rate 20%. In the independent data (squares), the agreement rate is at chance level (20%) across all lags. Dotted lines plot one standard deviation (over blocks) above and below the mean, so as to convey the overall variability between blocks.

Table 1: Overall performance for each observer, and the best-fitting parameter values (with α fixed at .01). In this table, ϕ denotes the motor-control parameter, while λ describes the degree to which participants integrated older trials. The λ_s parameter refers to the trials with sequential dependencies, and λ_i to the independent trials. Note that values of 0 or 10 denote the extreme values for the optimization, and both indicate the absence of interesting sequential dependencies. (See main text for details).

	sequential	independent	λ_s	λ_i	ϕ
MO	68.7%	60.1%	.286	.333	.878
AM	64.2%	59.8%	.030	.001	.593
MP	89.9%	86.8%	.153	.001	.136
BS	90.3%	91.9%	.148	.588	.140

ated the sequential dependencies by generating 1000 observations using a simple hidden Markov model that produced a slight bias towards more central locations: from left to right of screen, the marginal distribution over locations was [.143 .250 .246 .220 .122], which was broken in to 10 blocks of 100 trials. This produced the sequential structure shown by the circles in Figure 3. A second set of stimuli without any sequential structure were produced by randomly permuting the original items, and breaking the permuted set into 10 new blocks. These are illustrated with the squares in Figure 3.

Apparatus. Stimuli were presented on a 15in touchscreen (ELO Touch Systems, Elo Entuitive Systems). This is a standard CRT monitor overlaid with a touchsensitive layer. Pressure from the fingertip triggered a program that calculated the (x, y) position of the finger. The sensitivity of this information was tested by measuring the variability of the response to a regular artificial surface (an unused pencil eraser) touched with care to the same position. The error associated with this response was less than 0.1° in the x and y direction, and well within the error associated with a finger press. The experiment was conducted using custom written software from Matlab (Mathworks), using the Psychophysics Toolbox extensions (Brainard, 1997; Pelli, 1997).

Results

Overview. A brief summary of the data is presented in Table 1. As one might expect, the probability of successfully hitting the target tended to be higher when sequential dependencies were present in the data, but the effect was significant only for MO (Bayes factor[2] 31:1). For AM, MP and BS, the Bayes factors all favored a no-difference model (odds ratios 1:2.5, 1:7.7, 1:14 respectively). In the case of MP and BS this is not surprising, as both showed very high hit rates in both conditions, suggesting that they were probably not under a great deal of time pressure.

Modelling details. The specific version of the model that we fit to the data assumes that each observer has a single motor-correction parameter ϕ that is invariant across conditions, but may apply different intertemporal generalization rates depending on whether the data have sequential structure (λ_s) or are independent (λ_i). We present the results of the model fitting exercise in two ways. In order to provide a concise summary, Table 1 lists the best fitting parameter values for all four observers (Bayesian MAP estimators under uniform priors on $[0, 10]$ for all parameters).

Illustrative data. To provide a more concrete illustration of the raw data and the model predictions, Figure 4 plots the performance of MO during one representative block of trials containing sequential structure, and Figure 5 contains analogous plots for a block without such structure. The upper panels plot the predicted probability with which the observer would hit the target on every trial; white markers denote trials when MO succeeded in doing so, and black markers denote the miss trials. Lower panels plot the location of the target on every trial, again color-coded by the outcome of that trial. As is clear in both cases, the model tends to predict high hit-probability on the actual hits and low hit-probability on the misses, as one would hope if the model was performing well.

Model fit. Since space does not permit extensive quantitative evaluations, we simply examine how effectively the model is able to predict whether people were likely to hit the target. For each participant, we bin the trials according to the model predictions as to the likelihood of a hit, and then count the proportion of hits on those trials. As is shown in Figure 6, although the model fails in some respects, when the model predicted people were more likely to hit the target, they were in fact more likely to do so. Overall, the fits were slightly better in the sequential dependencies condition (white dots), but it is interesting that the model performs above chance even in the independent condition (black dots). That is, the model is only able to distinguish between trials on the basis of the locations of preceding trials, but nevertheless is able to discriminate between human hits and misses even when no actual sequential dependencies exist. This suggests that people are, in effect, "looking" for structure even when none exists.

In addition to the overall model fits presented in Figure 6, it is instructive to examine some key qualitative characteris-

[2]Model comparisons made using standard beta-binomial models (see Lee & Pope 2006)

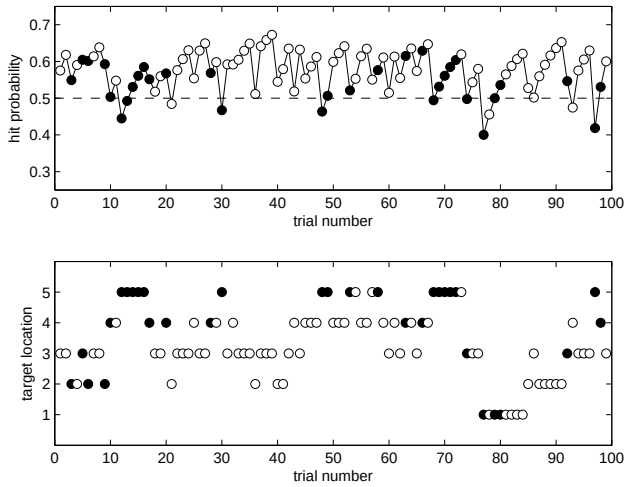

Figure 4: Performance of observer MO for a block containing sequential dependencies. White dots corresponds to trials in which the target was hit, while misses are shown with black dots. The lower panel shows the true location on each trial, while the upper panel shows the predicted probability of a hit according to the model.

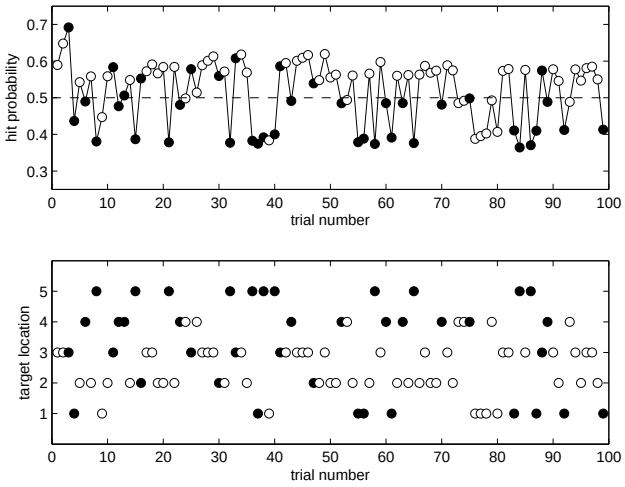

Figure 5: Performance of observer MO for a block without sequential dependencies. As before, white dots corresponds to trials in which the target was hit, while misses are shown with black dots. The lower panel shows the true location on each trial, while the upper panel shows the predicted probability of a hit according to the model.

tics of the data. One such effect relates to the length of the "run" preceding the current trial. By run length, we mean the number of trials that have elapsed since the last time the target was in a different location. So if the last five trials were in locations 2,1,4,4,4 we have a run length of 3. If the current trial is also in target location 4, then the run has continued. If it is in any other location, the run breaks. As is shown in Figure 7, the length of the run has an influence on the probability of hitting the target. The model successfully predicts that as the run length increases, the difference in hit probability between "run continuing" trials and "run break" trials gets larger. However, it does not always do so in precisely the right fashion (as is most clear for observer MO).

Discussion

Overall, it appears that observers do use the sequential information in the experimental trials to assist rapid reaching: all observers showed nonzero λ values for the sequential condition, suggesting they weighted recent information more heavily (only two participants did so in the independent condition). Additionally, one out of four observers, induces a significant improvement in hit rates.

When model performance was assessed, we found that it was able to predict performance reasonably well for all four observers, though there were some systematic deviations that suggest that further model development is needed. Moreover, when considered in terms of the key trend evident in data (effect of run length), the model was able to capture the basic patterns shown by the human participants, though not perfectly. This is encouraging because it suggests that despite its simplicity, the model has captured significant aspects of our data.

In terms of the experiment itself, it is important to note is that it was designed to have the second target presented late in the observer's reach. This was done to ensure that the observer would be under time pressure and would not have sufficient time to update their plan and therefore reveal, through their errors, what those priors were. Accordingly, the overall performance should have been kept to moderate levels. This manipulation was successful for two observers (AM and MO); however, since it did not succeed for MP and BS, it may be necessary to tailor the time of second target presentation individually for each observer.

General Discussion

Though this work is preliminary, the results of our experiment indicate that – consistent with what we expected given previous research – people do use prior history about target locations to update hand movements. In terms of the modelling exercise, it is encouraging to note that the simple model presented here provides a reasonably good account of individual participant behavior on a trial-to-trial basis, not merely in the aggregate. Furthermore, it also predicts specific qualitatively important trends, such as the change in change in hit probabilities that accompany different lengths of trial runs. With this in mind, future work will seek to extend the model to accommodate response biases (some locations may be easier to reach to than others) and to predict specific landing locations rather than just hit probabilities. In the meantime, however, this simple model represents a useful first approximation.

In terms of the broader research questions, one of the main goals in this line of work is to consider the extent to which people can integrate information from quite different sources and times scales. On the one hand, during a rapid movement the world is constantly changing both in an allocentric sense (the fly moves before you can swat it) and also an egocentric sense (eyes and hands move relative to the position of the body). On the other hand, in between these rapid movements people acquire rich conceptual representations of their world

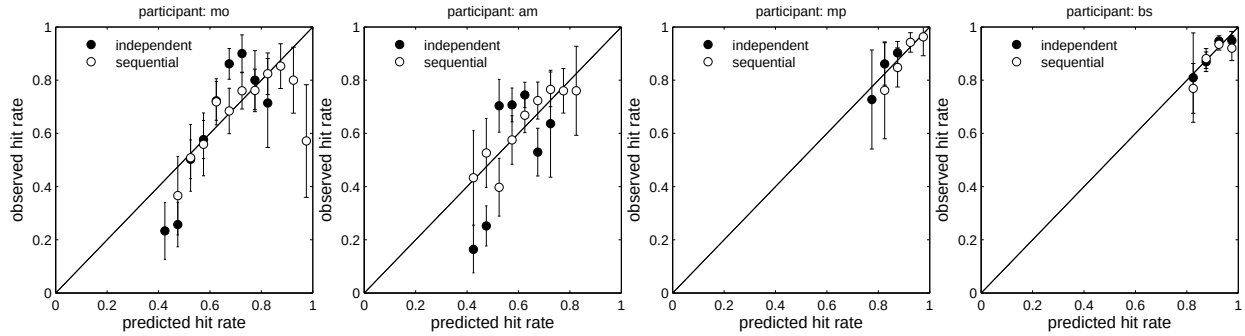

Figure 6: Plots illustrating the extent to which the model can discriminate between hits and misses for all four observers. Error bars show 95% confidence intervals, indicating that the model fails to capture some aspects to human performance. Nevertheless, the correlations between model predictions and human performance are reasonably high ($r = .72$ for MO, $r = .75$ for AM, $r = .91$ for MP and $r = .90$ for BS; all $p < .01$), indicating that the model is able to capture some of the variability in human responding.

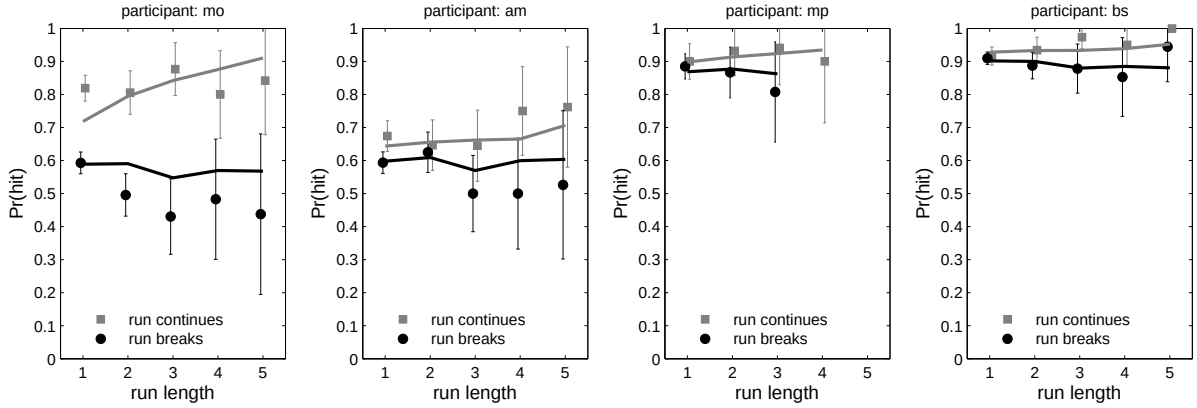

Figure 7: The effect of "run length" on performance, broken down by whether the run continues to the current trial (black markers) or whether the current trial appears in a different location to the previous ones (grey markers). Error bars show 95% confidence intervals. The model predictions (solid lines) are able to partially capture this pattern.

and how it is likely to change (e.g., flies move differently to mosquitos). It makes sense to think that people can integrate these sources of knowledge when needed, and somewhat reassuring to know that we do.

Acknowledgements

DJN was supported by an Australian Research Fellowship (ARC grant DP0773794). We thank the reviewers for helpful comments, and Riki Sutherland for helping to run the experiment.

References

Arulampalam, S., Maskell, S. & Gordon, N. (2002). A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. *IEEE Transactions on Signal Processing, 50*, 174-188.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision, 10*, 433-436.

Brown, S. D. & Steyvers, M. (2009). Detecting and predicting changes. *Cognitive Psychology, 58*, 49-67.

Fitts, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of the movement. *Journal of Experimental Psychology 47*, 381-391.

Gelman, A., Carlin, J.B., Stern, H.S. & Rubin, D.B. (1995). *Bayesian Data Analysis*. London: Chapman & Hall.

Jaynes, E. T. (2003). *Probability Theory: The Logic of Science*. Cambridge, UK: Cambridge University Press.

Körding, K. P. & Wolpert, D. (2004) Bayesian integration in sensorimotor learning. *Nature, 427*, 244-247.

Ma-Wyatt, A. & McKee, S. P. (2006). Visual uncertainty determines endpoint precision for rapid pointing. *Vision Research 46*, 4675-4683.

Ma-Wyatt, A. & McKee, S. P. (2007). Visual information is used continuously to determine endpoint precision for rapid pointing. *Experimental Brain Research, 179*, 55-64.

Medin, D.L. & Schafer, M. (1978). A context theory of classification learning. *Psychological Review, 85*, 207-238.

Lee, M.D. & Pope, K.J. (2006). Model selection for the rate problem: A comparison of significance testing, Bayesian, and minimum description length statistical inference. *Journal of Mathematical Psychology, 50*, 193-202

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

Reed, S.K. (1972). Pattern recognition and categorization. *Cognitive Psychology, 3*, 382-407.

Sober, S. J. & Sabes, P. N. (2005). Flexible strategies for sensory integration during motor planning. *Nature Neuroscience, 8*, 490-497.

Trommershäuser, J., Maloney, L. T. & Landy, M. S. (2003). Statistical decision theory and the selection of rapid, goal-directed movements. *Journal of the Optical Society of America: Image Sci. Vis. 20*, 1419-1433.

Yu, A. J. & Cohen, J. D. (2009). Sequential effects: Superstition or rational behavior? *Advances in Neural Information Processing Systems, 21*.