

Improved Probabilistic Inference as a General Learning Mechanism with Action Video Games

C. Shawn Green,^{1,2} Alexandre Pouget,¹ and Daphne Bavelier^{1,*}

¹Department of Brain and Cognitive Sciences, University of Rochester, Rochester, NY 14627, USA

Summary

Action video game play benefits performance in an array of sensory, perceptual, and attentional tasks that go well beyond the specifics of game play [1–9]. That a training regimen may induce improvements in so many different skills is notable because the majority of studies on training-induced learning report improvements on the trained task but limited transfer to other, even closely related, tasks ([10], but see also [11–13]). Here we ask whether improved probabilistic inference may explain such broad transfer. By using a visual perceptual decision making task [14, 15], the present study shows for the first time that action video game experience does indeed improve probabilistic inference. A neural model of this task [16] establishes how changing a single parameter, namely the strength of the connections between the neural layer providing the momentary evidence and the layer integrating the evidence over time, captures improvements in action-gamers behavior. These results were established in a visual, but also in a novel auditory, task, indicating generalization across modalities. Thus, improved probabilistic inference provides a general mechanism for why action video game playing enhances performance in a wide variety of tasks. In addition, this mechanism may serve as a signature of training regimens that are likely to produce transfer of learning.

Results

Visual Motion Direction Discrimination: Experiment 1

To test the hypothesis that video game experience leads to improved probabilistic inference, video game players (VGPs) and non-video game players (NVGPs) were first compared on a standard coherent dot motion direction discrimination task [14]. In this task, subjects viewed a dynamic random dot motion display and were asked to indicate the direction of coherent motion (left or right) as quickly and accurately as possible by means of a key press (Figure 1A).

One of the benefits of this task is that psychometric and chronometric curves are well captured by diffusion-to-bound models (DDM) or variations thereof [17–27] as well as a recent neural model of decision making [16], which, unlike the standard DDMs, has a clear probabilistic interpretation. This interpretation is based on recent theories in which neural patterns of activity represent probability distributions over the encoded variables [28–31]. Critically, this allows us to assess whether the changes that arise as a result of video game experience

meet our standards for improved probabilistic inference. These can be defined rigorously in the task we chose by considering decision making from a probabilistic perspective. Before committing to a choice, the best a subject can do is to compute the probability that each choice is correct given all the evidence up to the present time, a distribution also known as the posterior distribution over choices (which we denote $p(c|e)$ where c are the choices and e is the evidence). This computation requires knowledge of $p(e|c)$, the statistics of the evidence. Indeed, through Bayes rule, $p(c|e)$ is proportional to $p(e|c)$. In general, subjects do not initially know the statistics of the evidence, which in turn means that the calculated posterior distribution over choices is not the optimal one. However, through repeated exposure to a task, subjects have the opportunity to learn these statistics and, as a result, can make decisions on the basis of a more accurate posterior distribution. This type of improved inference is what we mean by making better use of the evidence.

The DDM fits to the data provide support for this hypothesis. The results show an overall decrease in decision bound, no change in nondecision time, and crucially a greater integration rate in the VGP population (Figures 1B and 1C). This latter result indicates increased sensitivity in VGPs suggesting that VGPs may indeed make better use of the evidence than NVGPs. Yet to definitively make this point, it is necessary to demonstrate that VGPs use better posterior distributions over choices than those used by NVGPs. Because the DDM used here does not provide a natural way to compute a posterior distribution (particularly when the coherence varies over trials), we took advantage of a recent neural model of probabilistic decision making [16] in which the state of the network encodes the posterior distribution over choices given the evidence (Figure 2A).

We first used this model to fit the psychometric and chronometric curves of VGPs and NVGPs via a numerical optimization procedure (Figure 2B; for details see [Supplemental Experimental Procedures](#), part A available online). Data from VGPs were well modeled by assuming only an increase in the strength of the feedforward connectivity between the two layers of the network representing areas MT and LIP as compared to NVGPs. A change in this single parameter naturally accounts for the finding that only RTs change whereas accuracy stays constant. Consistent with this proposal, Law and Gold [32, 33] recently showed that improved performance on a motion discrimination task after extensive training on the motion task itself is also captured by strengthening the connections between the sensory and integration layers.

We then asked whether action video game experience is indeed associated with a better posterior distribution for a fixed amount of evidence. In the case of a binary decision task (i.e., left versus right), the quality of this posterior distribution can be assessed by calculating the log odds (i.e., the log of the ratio of the probability that the dots move rightward over the probability that the dots move leftward). Because the log odds in the model reflects the quality of the evidence available to the decision maker, if action video-game players do make better use of the evidence, we should find an increase in the absolute value of the log odds in their model. This is indeed

*Correspondence: daphne@cvs.rochester.edu

²Present address: Department of Psychology, Center for Cognitive Sciences, University of Minnesota, Minneapolis, MN 55455, USA

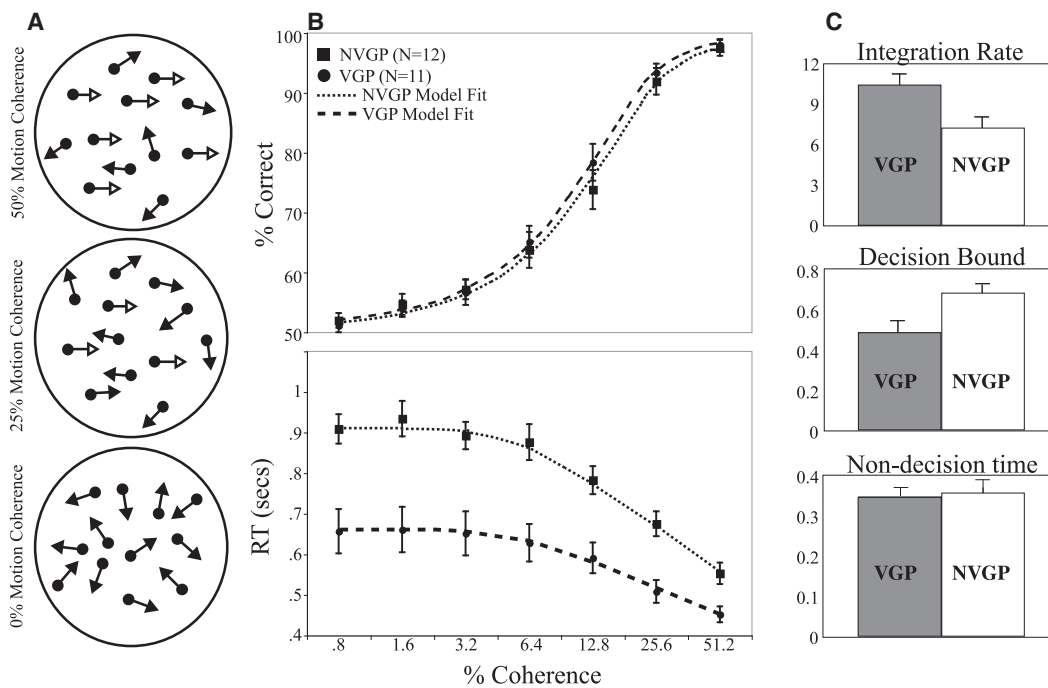


Figure 1. Visual Motion Direction Discrimination

(A) Task. Subjects viewed a dynamic random dot motion display and were asked to indicate the net direction of motion (left or right; here, the correct answer would be right). On every trial, some proportion of the dots moved coherently (top, 50% coherence; middle, 25% coherence; bottom, 0% coherence) either to the left or to the right, while the remaining dots were replotted randomly. By parametrically varying the number of coherently moving dots from very few to many, full psychometric and chronometric curves could be obtained.

(B) Behavior. Although VGPs and NVGPs demonstrated equivalent accuracy ($p = .65$, $p\text{-}\eta^2 = .01$) (top), VGPs responded substantially faster than NVGPs ($F(1,21) = 18.9$, $p < .001$, $p\text{-}\eta^2 = .47$) (bottom). Importantly, VGP status interacted with coherence because of a greater decrease in RTs in VGPs at low than high coherence ($F(6, 126) = 3.5$, $p < .001$, $p\text{-}\eta^2 = .15$).

In this and all other psychometric and chronometric curve figures, error bars correspond to between-subject standard error.

(C) Drift diffusion model (DDM). The accumulation of the noisy sensory evidence is simulated by the diffusion of a particle upward or downward until a decision bound is reached. DDM models generate psychometric and chronometric curves that are constrained by three main variables [14]: (1) the rate at which information is accumulated over time, (2) the height of the decision bound at which the system stops accumulating evidence and a decision is made, and (3) the nondecision time, an additive amount of time that is common to all tasks and reflects nondecision processes such as motor planning and execution. To quantitatively assess the individual contribution of integration rate, decision bound, and nondecision time, RT and accuracy data were simultaneously fit to each subject's data with the proportional-rate diffusion model as in Palmer and colleagues [14]. The fits were good and equivalent in the two groups ($r^2_{\text{VGP}} = .93$, $r^2_{\text{NVGP}} = .90$, $p = .36$). The rate of integration was greater in the VGP than the NVGP group ($t(21) = 2.6$, $p = .02$, Cohen's $d = 1.13$, top), whereas the opposite result was observed for the decision bound ($t(21) = 3.6$, $p = .002$, Cohen's $d = 1.57$, middle). No difference was observed between the groups in nondecision time ($p > .7$, Cohen's $d = 0.14$, bottom), eliminating an additive, postdecision process as a possible source of group differences. Data from individual subjects were fit separately and error bars correspond to between-subject standard error.

what was observed, with the log odds being 19% higher on average across all coherences, confirming that, according to the probabilistic population network, VGPs make better use of the evidence than NVGPs (for details see [Supplemental Experimental Procedures](#), part A).

Auditory Tone Location Discrimination: Experiment 2

A key question concerns the generalization of this finding to other tasks and domains. To ask whether the improved probabilistic inference noted in VGPs generalizes to other tasks and domains, an auditory analog of the motion direction task was developed (Figure 3A, for details see [Supplemental Experimental Procedures](#), part B). As in the motion task, performance on the task requires the accumulation of information over time, and thus the intuitions regarding sensitivity and decision bound are similar and can be modeled in the same manner. The results in Experiment 2 mirror those of Experiment 1 quite closely (Figure 3B). In the DDM, the integration rate parameter was significantly greater in VGPs than in NVGPs, confirming greater sensitivity to the stimulus in VGPs. The decision

bound parameter was significantly smaller in VGPs than in NVGPs and no significant difference was observed in the nondecision time parameter. The neural model confirmed that the difference between NVGPs and VGPs can be captured by changing solely the strength of the feed-forward connections (by 51%, Figure S1A). We further confirmed that VGPs do indeed make a better use of the evidence in this paradigm, with log odds being 48% higher on average across all levels of signal to noise ratio in VGPs as compared to NVGPs.

Video Game Training: Experiment 3

Although our hypothesis is that extensive video game experience is at the root of these changes, it could also be the case that VGPs are individuals who have been born with improved abilities at performing probabilistic inferences. To establish that video game experience is sufficient to drive a better use of the evidence, NVGPs underwent 50 hr of video game training on either action video games or a control video game. If action video game experience does improve probabilistic inferences in neural circuits, larger improvements should

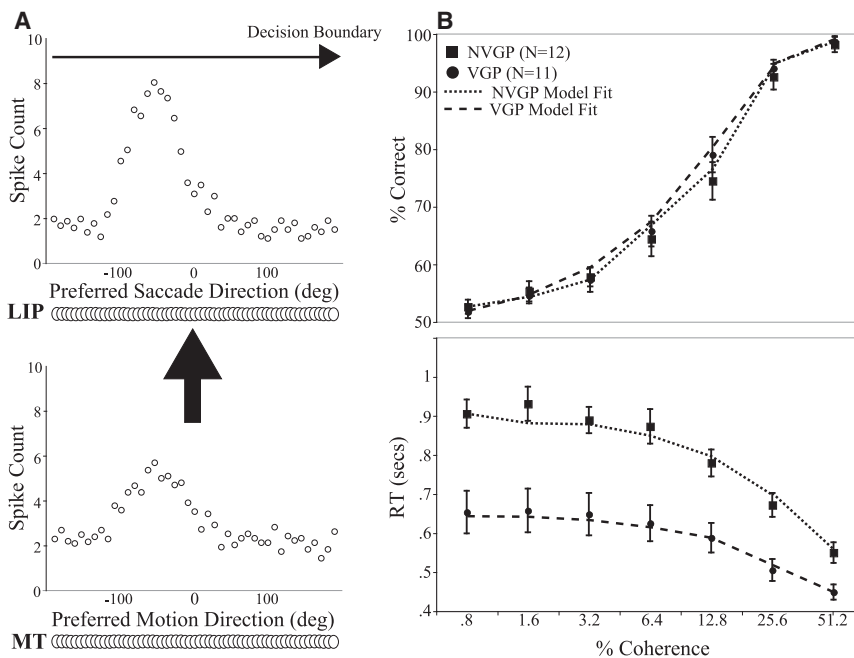


Figure 2. Neural Model—Motion Direction Discrimination

(A) Neural Network Architecture. The network consists of two interconnected layers of neurons with Gaussian tuning curves. In MT, the sensory layer, the tuning curves are for direction of motion, whereas in LIP, the integration layer, the tuning curves are for saccade direction, as a proxy in our case for a left/right decision. Note that we do not mean to say that LIP is the only area involved in this process—the label is used mostly for convenience (the same is true for the MT label in the input layer). The layers differ by their connectivity and dynamics. The MT neurons send feed-forward connection to the LIP neurons. Each LIP neuron receives a Gaussian pattern of weight centered on the MT neuron with the same preferred direction. The LIP neurons also have lateral connections to implement short-range excitation and long-range inhibition as well as a long time constant (1 s) allowing them to integrate their input. Each panel indicates a representative pattern of activity in terms of spike count 200 ms into a trial for the sensory layer (MT, bottom) and the integration layer (LIP, top).

(B) Neural Model Fit. As with the DDM, the fits were good and equivalent for the two groups. The neural model captures the change in performance from NVGP to VGP with a 55% increase

in the conductance of the feed-forward connections between the sensory (MT) and the integration (LIP) layers and, in contrast to the DDM, nearly no change in the bound height (a 1% decrease, which is within the resolution of our numerical maximization procedure; see [Supplemental Experimental Procedures](#), part A). The conductance controls the amount of information transmitted from the sensory layer to the integration layer per unit time, or the strength of the feed-forward connections. It is important to note that this effect is not analogous to a simple change in sensitivity in DDMs. Although increasing the conductance does increase the amount of information transmitted from the sensory to the integration layer (which intuitively should increase accuracy), it also induces large fluctuations in the membrane potential of the neurons. These fluctuations lead the network to reach the bound faster, thus lowering the percentage of correct responses. These two effects cancel one another over a wide range of parameters, allowing a single change in feed-forward strength to alter reaction time while leaving the psychometric curve nearly unchanged. Model fit corresponds to best fit to the mean data rather than the mean of individual fits.

be noted in the action-trained group than in the control group. This prediction was confirmed because the results of this training study are nearly identical to the results of Experiments 1 and 2 ([Figure S2](#), for details see [Supplemental Experimental Procedures](#), part C). Critically, better use of the evidence was noted only in the action group at posttest. The log odds for the motion task were 16% (respectively 31% for auditory) higher on average across all levels of coherence (respectively SNR) at posttest as compared to pretest for the action group (see [Figure S3](#)). The results of this experiment therefore confirm the causal relationship between action video game experience and improvements in probabilistic inference.

Ruling out Alternative Accounts—Critical Duration Study: Experiment 4

A remarkable feature of our results is that action video game experience leads to a decrease in RTs while leaving accuracy unchanged. At first sight, this pattern could be consistent with a reduction in nondecision time, for instance, through a decrease in motor latencies. However, a change in nondecision time alone cannot capture the present data because it predicts a constant difference between VGP and NVGP reaction time as coherence varies, whereas we report larger reaction times differences at lower than higher signal strengths. This was confirmed by our fitting procedure, which revealed that the sensitivity and the bound height must be changed to capture our data, and that nondecision time is in fact not significantly different across populations. Nonetheless, to confirm our key result, namely that the sensitivity to stimuli

has increased between NVGPs and VGPs, we ran our motion and auditory tasks in a condition in which stimuli were presented for fixed durations and the subjects were instructed to watch/listen to the entire stimulus before reporting their decision. We then plotted the accuracy of the response as a function of duration. Because this task does not involve speeded reaction time, nondecision time is irrelevant in accounting for these data (for details see [Supplemental Experimental Procedures](#), part D).

As predicted by higher sensitivity in VGPs than in NVGPs, we found the rate with which accuracy increases as a function of duration (the parameter β in [Figure 4](#)) to be higher in VGPs [34, 35]. Simulations of the drift diffusion model with fixed duration confirmed that a change in sensitivity between VGPs and NVGPs is required to account for the data in [Figure 4](#). We first ran a simulation in which the sensitivity for VGPs and NVGPs were set to be equal, while the bounds heights were independently adjusted for VGPs and NVGPs to maximize the fit to the experimental data. This simulation revealed that it is impossible to capture the faster rise of accuracy for VGPs, particularly for short durations ([Figures S4A](#) and [S4B](#)). Next, to determine whether the change in sensitivity estimated from Experiments 1 and 2 can account for the critical duration data, we ran simulations in which the sensitivity ratios between VGPs and NVGPs were the same as estimated from the reaction time experiments (E1 and E2). Under these conditions, the model captures the results in [Figure 4](#) well, thus confirming that sensitivity must be higher in VGPs compared to NVGPs ([Figures S4C](#) and [S4D](#)).

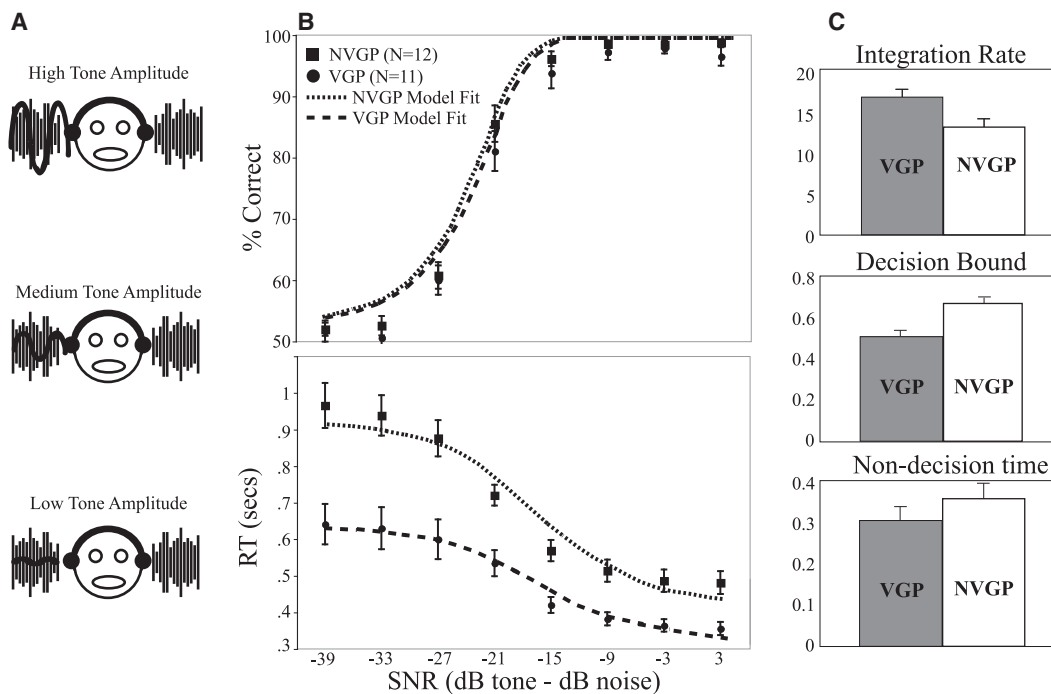


Figure 3. Auditory Tone Location Discrimination

(A) Task. A pure tone embedded in a white noise mask was presented in one ear, while white noise alone was presented in the other (both ears being normalized to the same mean amplitude). The subjects' task was to indicate with a button press the ear in which the tone was present as quickly and accurately as possible. In a manner consistent with adjusting the coherence level of the motion stimulus, the ratio of the amplitude of the target tone to the white noise mask was manipulated in order to test performance across the range of possible accuracy levels and reaction times (high amplitude, top; low amplitude, bottom).

(B) Behavior. As in Experiment 1, VGPs and NVGPs demonstrated equivalent accuracy ($p = .32$, $p\text{-}\eta^2 = .05$) (top), VGPs responded substantially faster than the NVGPs ($F(1,21) = 20.6$, $p < .001$, $p\text{-}\eta^2 = .50$) (bottom), and the RT difference between groups was greater at lower signal-to-noise ratios (SNR) ($F(7,147) = 5.2$, $p < .001$, $p\text{-}\eta^2 = .2$).

(C) Drift diffusion model. The rate of integration was greater in the VGP than the NVGP group ($t(21) = 3.8$, $p = .001$, Cohen's $d = 1.66$) (top), while the opposite result was observed for the decision bound ($t(21) = 2.6$, $p = .02$, Cohen's $d = 1.13$) (middle). No difference was observed between the groups in nondecision time ($p > .05$, Cohen's $d = .82$) (bottom). Data from individual subjects were fit separately and error bars correspond to between-subject standard error.

Discussion

Action video game experience results in more efficient use of sensory evidence. Importantly, these improvements are not restricted to the visual modality, but appear in the auditory modality as well. Moreover, 50 hr of action video game training led to qualitatively similar results in a group of NVGPs, establishing a causative relationship between action video game experience and improved probabilistic inference.

To establish the role of improved statistical inference in VGP performance, full chronometric and psychometric curves were measured. Using DDMs, increased sensitivity (along with a compensatory decrease in bound) provided the best fit to the data, thus confirming the hypothesis that video game play results in a more efficient use of the evidence. That statistical inference is improved in gamers was further confirmed by establishing enhanced performance in gamers in fixed duration paradigms. Note that these results were far from being a foregone conclusion. Although faster RTs in gamers were expected, this could have resulted from a speed-accuracy tradeoff through a change in the decision bound and/or faster motor execution through a change in the nondecision time, without concomitant changes in sensitivity. In contrast, the six experiments presented establish the role of increased sensitivity in gamers' performance.

The neural model framework provides converging evidence for this claim because it captures VGP behavior by enhancing the connection strength between the layer providing the momentary sensory evidence and the layer integrating the evidence, leading to a more efficient use of the evidence. Increasing the connection strength not only increases the amount of information per unit of time transmitted to the output layer but also induces larger fluctuations in the membrane potential of neurons, which in turn lead the network to reach the bound faster. The net effect on accuracy is negligible because the output layer receives more information per unit of time, but integrates information over shorter durations. As a result, the percentage of correct responses at the bound remains stable. A novel contribution from the neural model therefore is that it naturally accounts for one striking regularity of the data with just one parameter change, and crucially this parameter controls the quality of the statistical inference. One might wonder why the nervous system does not keep the feedforward weights set to a maximum value since our results show that increasing the feedforward weights leads to better performance. This might be in part because of the metabolic cost of maintaining synapses with high efficacies. But, more importantly, it is not the case that increasing the feedforward weights always increases performance. We have found analytically that there is an optimal value of the

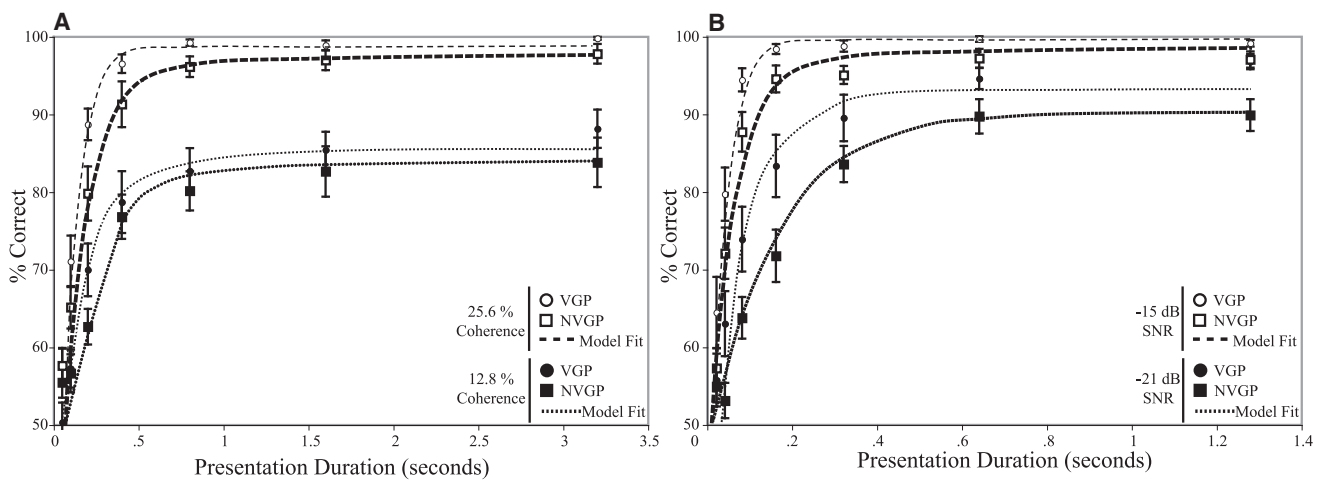


Figure 4. Critical Duration Experiment Results

Accuracy of the VGPs and NVGPs for two levels of visual coherence (A) and two levels of auditory SNRs (B) as a function of presentation duration (see Supplemental Information).

Individual subject data were modeled as a simple exponential rise to an asymptote (VGP = thinner lines; NVGP = thicker lines) using $\%Correct(t) = \lambda(1 - e^{-\beta(t - \delta)}) + 50\%$, where λ is the level of asymptotic performance, β is the rate at which accuracy grows as a function of time, and δ is the intercept or the time at which accuracy rises above chance levels [35]. In both the DDM and the neural model, faster accumulation of information predicts greater rate (β) value in the VGPs. This prediction was confirmed (visual motion: VGP, 8.1 ± 1.1 ; NVGP, 4.8 ± 1.1 ; $F(1,21) = 4.6$, $p = .045$, $p\text{-}\eta^2 = .19$; auditory tone: VGP, 24.5 ± 3.9 ; NVGP, 12.9 ± 3.8 ; $F(1,21) = 4.5$, $p = .044$, $p\text{-}\eta^2 = .18$), demonstrating again a greater rate of sensory integration in VGPs.

feedforward weights, beyond which a further increase worsens performance (J. Beck, V. Bejjanki, and A.P., unpublished data). The optimal value of the weights depends on the statistics of the spike trains in the sensory layer.

The fact that action gamers perform better probabilistic inference is appealing because it provides a unified mechanism to explain why action video game training improves participants' performance in seemingly widely different tasks such as contrast detection, visual search, multiple object tracking, letter recognition with flankers, and decision making [1, 6, 7, 9, 36]. However, such an improvement may also have much wider implications. Decision making can be formalized as a probabilistic inference in which subjects must compute the posterior distribution over choices given the evidence at each time step [37–39]. This iterative process requires that the posterior distribution at the previous time step be multiplied by the likelihood of the evidence at the present time step (see Equation A1 in Supplemental Experimental Procedures). To be done optimally (i.e., without loss of information), this process requires both that the multiplication step be faithfully executed and also that the subjects have perfect knowledge of the likelihood function, which is to say, perfect knowledge of the statistics of their sensory input. It is highly unlikely that our subjects came to the task equipped to perform optimally from the get-go. At the beginning of the experiment, our subjects were naive to the visual and auditory stimuli. Their lack of knowledge of the statistics of our particular stimuli would initially result in a poor estimate of the likelihood of the visual or auditory evidence. Similarly, subjects were initially naive to the inference task used in our studies. Consequently, they had to learn the statistics of our particular stimulus set (more accurate likelihood of the evidence) as well as the appropriate inference for this task (performing the multiplication that permits the update over time of the posterior distribution). One possibility is that VGPs perform better in those tasks because they learn a better model of the stimuli

used and/or perform the inference more accurately [32], which in turn leads to a better posterior distribution.

This proposal is closely related to a dominant idea in the field of perceptual learning by which learning occurs through template matching by reweighting the connectivity between sensory and decision stages [40–42]. Law and Gold [32, 33] have recently documented how the effective connectivity between MT and LIP changes as monkeys are trained and then tested with motion stimuli similar to ours. Our results build on this work but establish that such a reweighting is not necessarily the product of extensive training with a specific class of stimuli as is standard in the field of perceptual learning. Indeed, improvements after action game training are not limited to playing the game itself, but generalize to new tasks. Gamers, we propose, acquire the ability to dynamically retune the connectivity between the momentary evidence layer and the layer integrating the evidence based on the statistics of the very task they are performing.

This type of learning may be a consequence of the nature of action video game training. Unlike standard learning paradigms, which have a highly specific solution, there is no such specific solution in action video games because situations are rarely, if ever, repeated. Thus, the only characteristics that can be learned are how to rapidly and accurately learn the statistics on the fly and how to accumulate this evidence more efficiently. We can only speculate as to what might be the neural mechanisms responsible for this remarkable transfer of learning, because there are several possibilities. First, the neural circuits involved in evidence accumulation might be shared across modalities, a plausible explanation given that the cortical areas that have been implicated in evidence accumulation are often multimodal [43]. Second, these areas are likely to be under the influence of shared fronto-parietal networks, providing a common source for performance improvements, a view supported by the impact of working memory training in cognitive improvements [44–46].

Finally, learning might involve the global release of neuromodulators that improve probabilistic inferences across all circuits, as has been proposed for instance for noradrenalin [47]. An important issue for future work will be to determine the relative role of these mechanisms in fostering such a wide transfer of learning.

Supplemental Information

Supplemental Information includes Supplemental Experimental Procedures and four figures and can be found with this article online at [doi:10.1016/j.cub.2010.07.040](https://doi.org/10.1016/j.cub.2010.07.040).

Acknowledgments

We thank Ashley Anderson, Renjie Li, and David Feder for help with subjects, Mike Sugarman for help with figure preparation, Jan Drugowitsch for help with DDM simulations, and Josh Tenenbaum for helpful discussions. This work was supported by National Institutes of Health grant EY016880 and Office of Naval Research grant N00014-07-1-0937.3 to D.B., NIDA grant #DA022780 and a research grant from the James S. McDonnell Foundation to A.P., as well as award P30 EY001319 from the National Eye Institute. The content of the manuscript is solely the responsibility of the authors and does not necessarily represent the official views of the National Eye Institute or the National Institutes of Health.

Received: May 4, 2010

Revised: June 29, 2010

Accepted: July 27, 2010

Published online: August 26, 2010

References

- Castel, A.D., Pratt, J., and Drummond, E. (2005). The effects of action video game experience on the time course of inhibition of return and the efficiency of visual search. *Acta Psychol. (Amst.)* 119, 217–230.
- Greenfield, P.M., DeWinstanley, P., Kilpatrick, H., and Kaye, D. (1994). Action video games and informal education: Effects on strategies for dividing visual attention. *J. Appl. Dev. Psychol.* 15, 105–123.
- Bialystok, E. (2006). Effect of bilingualism and computer video game experience on the Simon task. *Can. J. Exp. Psychol.* 60, 68–79.
- Orosy-Fildes, C., and Allan, R.W. (1989). Psychology of computer use: XII. Videogame play: Human reaction time to visual stimuli. *Percept. Mot. Skills* 69, 243–247.
- Feng, J., Spence, I., and Pratt, J. (2007). Playing an action video game reduces gender differences in spatial cognition. *Psychol. Sci.* 18, 850–855.
- Green, C.S., and Bavelier, D. (2003). Action video game modifies visual selective attention. *Nature* 423, 534–537.
- Green, C.S., and Bavelier, D. (2006). Enumeration versus multiple object tracking: The case of action video game players. *Cognition* 101, 217–245.
- Green, C.S., and Bavelier, D. (2006). Effect of action video games on the spatial distribution of visuospatial attention. *J. Exp. Psychol. Hum. Percept. Perform.* 32, 1465–1478.
- Green, C.S., and Bavelier, D. (2007). Action-video-game experience alters the spatial resolution of vision. *Psychol. Sci.* 18, 88–94.
- Fahle, M. (2005). Perceptual learning: Specificity versus generalization. *Curr. Opin. Neurobiol.* 15, 154–160.
- Liu, Z., and Weinschall, D. (2000). Mechanisms of generalization in perceptual learning. *Vision Res.* 40, 97–109.
- Jeter, P.E., Doshier, B.A., Petrov, A., and Lu, Z.L. (2009). Task precision at transfer determines specificity of perceptual learning. *J. Vis.* 9, 1–13.
- Xiao, L.Q., Zhang, J.Y., Wang, R., Klein, S.A., Levi, D.M., and Yu, C. (2008). Complete transfer of perceptual learning across retinal locations enabled by double training. *Curr. Biol.* 18, 1922–1926.
- Palmer, J., Huk, A.C., and Shadlen, M.N. (2005). The effect of stimulus strength on the speed and accuracy of a perceptual decision. *J. Vis.* 5, 376–404.
- Roitman, J.D., and Shadlen, M.N. (2002). Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task. *J. Neurosci.* 22, 9475–9489.
- Beck, J.M., Ma, W.J., Kiani, R., Hanks, T., Churchland, A.K., Roitman, J.D., Shadlen, M.N., Latham, P.E., and Pouget, A. (2008). Probabilistic population codes for Bayesian decision making. *Neuron* 60, 1142–1152.
- Ratcliff, R., and McKoon, G. (2008). The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Comput.* 20, 873–922.
- Ditterich, J., Mazurek, M.E., and Shadlen, M.N. (2003). Microstimulation of visual cortex affects the speed of perceptual decisions. *Nat. Neurosci.* 6, 891–898.
- Laming, D.R.J. (1968). *Information Theory of Choice-Reaction Times* (London: Academic Press).
- Link, S.W. (1992). *The Wave Theory of Difference and Similarity* (Hillsdale, NJ: Lawrence Erlbaum Associates).
- Mazurek, M.E., Roitman, J.D., Ditterich, J., and Shadlen, M.N. (2003). A role for neural integrators in perceptual decision making. *Cereb. Cortex* 13, 1257–1269.
- Ratcliff, R., and Rouder, J.N. (1998). Modeling response times for two-choice decisions. *Psychol. Sci.* 9, 347–356.
- Smith, P.L., and Ratcliff, R. (2004). Psychology and neurobiology of simple decisions. *Trends Neurosci.* 27, 161–168.
- Stone, M. (1960). Models for choice-reaction time. *Psychometrika* 25, 251–260.
- Usher, M., and McClelland, J.L. (2001). The time course of perceptual choice: The leaky, competing accumulator model. *Psychol. Rev.* 108, 550–592.
- Wald, A. (1947). *Sequential Analysis* (New York: Wiley).
- Wald, A., and Wolfowitz, J. (1947). Optimum character of the sequential probability ratio test. *Ann. Math. Stat.* 19, 326–339.
- Ma, W.J., Beck, J.M., Latham, P.E., and Pouget, A. (2006). Bayesian inference with probabilistic population codes. *Nat. Neurosci.* 9, 1432–1438.
- Jazayeri, M., and Movshon, J.A. (2006). Optimal representation of sensory information by neural populations. *Nat. Neurosci.* 9, 690–696.
- Deneve, S., Latham, P.E., and Pouget, A. (2001). Efficient computation and cue integration with noisy population codes. *Nat. Neurosci.* 4, 826–831.
- Huys, Q.J., Zemel, R.S., Natarajan, R., and Dayan, P. (2007). Fast population coding. *Neural Comput.* 19, 404–441.
- Law, C.-T., and Gold, J.I. (2008). Neural correlates of perceptual learning in a sensory-motor, but not a sensory, cortical area. *Nat. Neurosci.* 11, 505–513.
- Law, C.-T., and Gold, J.I. (2009). Reinforcement learning can account for associative and perceptual learning on a visual-decision task. *Nat. Neurosci.* 12, 655–663.
- Carrasco, M., McElree, B., Denisova, K., and Giordano, A.M. (2003). Speed of visual processing increases with eccentricity. *Nat. Neurosci.* 6, 699–701.
- Doshier, B.A. (1979). Empirical approaches to information processing: Speed-accuracy tradeoff functions or reaction time—A reply. *Acta Psychol. (Amst.)* 43, 347–359.
- Li, R., Polat, U., Makous, W., and Bavelier, D. (2009). Enhancing the contrast sensitivity function through action video game training. *Nat. Neurosci.* 12, 549–551.
- Yu, A.J., Dayan, P., and Cohen, J.D. (2009). Dynamics of attentional selection under conflict: toward a rational Bayesian account. *J. Exp. Psychol. Hum. Percept. Perform.* 35, 700–717.
- Vul, E., Frank, M.C., Alvarez, G.A., and Tenenbaum, J.B. (2009). Explaining human multiple object tracking as resource-constrained approximate inference in a dynamic probabilistic model. *Adv. Neural Inf. Process. Syst.* 22, 1955–1964.
- Ma, W.J., and Huang, W. (2009). No capacity limit in attentional tracking: Evidence for probabilistic inference under a resource constraint. *J. Vis.* 9, 3.1–3.30.
- Li, R.W., Levi, D.M., and Klein, S.A. (2004). Perceptual learning improves efficiency by re-tuning the decision ‘template’ for position discrimination. *Nat. Neurosci.* 7, 178–183.
- Petrov, A.A., Doshier, B.A., and Lu, Z.L. (2005). The dynamics of perceptual learning: an incremental reweighting model. *Psychol. Rev.* 112, 715–743.
- Jacobs, R.A. (2009). Adaptive precision pooling of model neuron activities predicts the efficiency of human visual learning. *J. Vis.* 9, 22.1–22.15.
- Bremmer, F., Schlack, A., Shah, N.J., Zafiris, O., Kubischik, M., Hoffmann, K., Zilles, K., and Fink, G.R. (2001). Polymodal motion

processing in posterior parietal and premotor cortex: A human fMRI study strongly implies equivalencies between humans and monkeys. *Neuron* 29, 287–296.

44. Jaeggi, S.M., Buschkuhl, M., Jonides, J., and Perrig, W.J. (2008). Improving fluid intelligence with training on working memory. *Proc. Natl. Acad. Sci. USA* 105, 6829–6833.
45. Persson, J., and Reuter-Lorenz, P.A. (2008). Gaining control: training executive function and far transfer of the ability to resolve interference. *Psychol. Sci.* 19, 881–888.
46. Light, K.R., Kolata, S., Wass, C., Denman-Brice, A., Zagalsky, R., and Matzel, L.D. (2010). Working memory training promotes general cognitive abilities in genetically heterogeneous mice. *Curr. Biol.* 20, 777–782.
47. Servan-Schreiber, D., Printz, H., and Cohen, J.D. (1990). A network model of catecholamine effects: Gain, signal-to-noise ratio, and behavior. *Science* 249, 892–895.