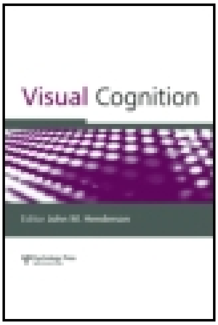

Visual Cognition
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/pvis20

Feedback-driven tuning of statistical summary representations

Judith E. Fan[a], Nicholas B. Turk-Browne[a] & Jordan A. Taylor[a]
[a] Department of Psychology, Princeton University, Princeton, NJ, USA
Published online: 10 Oct 2013.

To cite this article: Judith E. Fan, Nicholas B. Turk-Browne & Jordan A. Taylor , Visual Cognition (2013): Feedback-driven tuning of statistical summary representations, Visual Cognition

To link to this article: http://dx.doi.org/10.1080/13506285.2013.844961

Visual Cognition, 2013
http://dx.doi.org/10.1080/13506285.2013.844961

Routledge
Taylor & Francis Group

Feedback-driven tuning of statistical summary representations

Judith E. Fan, Nicholas B. Turk-Browne, and Jordan A. Taylor

Department of Psychology, Princeton University, Princeton, NJ, USA

The amount of sensory data encountered by the visual system often exceeds its processing capacity. One solution is to exploit statistical structure in the natural environment to generate a more efficient representation of the information (Simoncelli & Olshausen, 2001). For example, the visual system may construct a "statistical summary representation" over groups of visual objects, reflecting their general properties (Alvarez, 2011). Indeed, it has been shown that observers are able to quickly and accurately extract average values over a range of visual feature dimensions, including size (Chong & Treisman, 2003), orientation (Parkes, Lund, Angelucci, Solomon, & Morgan, 2001), and emotional expression (Haberman & Whitney, 2007). However, it remains an open question as to how observers learn to produce such accurate estimates of these summary statistics. Although good performance on these tasks suggests that summary features are readily accessible, it is not clear to what extent these statistical operations are performed automatically—integrating over sensory information in an unsupervised fashion, or are penetrable to task demands—flexibly incorporating observer goals and error-related feedback to maximize performance (Bauer, 2009; Myczek & Simons, 2008).

In the present study, we sought to understand the role of learning in statistical summary representations. Specifically, we examined the contribution of task practice and performance feedback to perceptual discrimination of the centroid (i.e., mean location) of a set of objects (Alvarez & Oliva, 2008). We hypothesized

Please address all correspondence to Judith Fan, Green Hall, Department of Psychology, Princeton University, Princeton NJ 08540, USA. E-mail: jefan@princeton.edu

This article reports preliminary findings from ongoing experiments and constitutes the proceedings of a talk given at the 21st Annual Meeting on Object Perception, Attention, and Memory. This work was supported by NINDS R01NS084948 (J.A.T.) and a NSF GRF (J.E.F).

that providing vector error feedback (i.e., containing both distance and direction information) while observers practised making pointing movements towards the centroid would improve the fidelity of their centroid representations. This improvement might be reflected in reduced error during training and lower discrimination thresholds in an independent perceptual test.

METHODS

The experimental protocol consisted of three phases conducted over two consecutive days: a pretest phase on Day 1, and training and posttest phases on Day 2. In all phases, trials contained an array of eight dots (diameter = 0.9°) presented for 200 ms on a touchscreen display (49.9° × 40.0°), positioned vertically 43 cm from the observer (Figure 1A). Individual dot locations were independently sampled from a bivariate Gaussian distribution (circularly symmetric; σ of marginal distributions = 11.2°).

The pretest and posttest phases consisted of a perceptual discrimination task (four blocks, 50 trials/block). This task entailed judging which of the four display

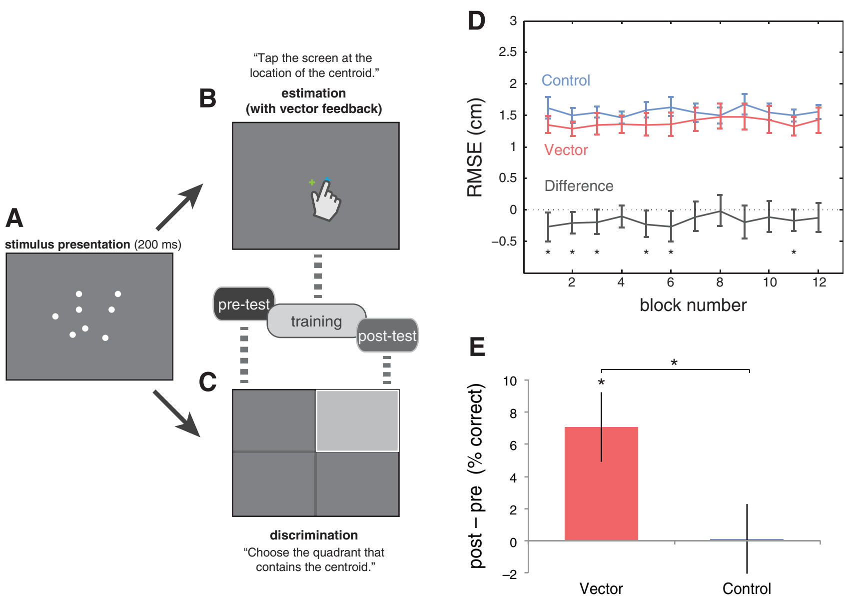

Figure 1. (A) Dot array presented on each trial for 200 ms. (B) Estimation task entailed tapping screen at location of centroid (training phase). Vector feedback (for Vector group) delivered as green crosshair at centroid location. (C) Discrimination task entailed selecting quadrant containing centroid (pretest and posttest phases). (D) Estimation performance during the training phase (error bars represent 95% CI). (E) Discrimination performance during the pretest and posttest phases (error bars represent 1 SEM). To view this figure in colour, please see the online issue of the Journal.

quadrants (NE, NW, SE, SW) contained the centroid (Figure 1C). Centroid locations in the discrimination task were distributed following a hyperbolic function:

$$r = c/\text{sqrt}(\cos(\theta)^{*}\sin(\theta)),$$

where r is the radial distance from the centre of the display, θ represents randomly sampled angles, and c is a scaling factor that controls the overall distance to the nearest quadrant boundaries. Smaller values of c produce displays with centroids falling closer to boundaries, thus making discrimination more difficult. The difficulty of the pretest and posttest phases (determined by the magnitude of c) was adjusted for each observer via an adaptive staircasing procedure calibrated to estimate their 62.5% threshold on the 4AFC task (QUEST; Watson & Pelli, 1983). No feedback was given during the pretest and posttest phases.

The training phase consisted of a centroid estimation task (12 blocks, 50 trials/block). This task entailed tapping the display at the perceived location of the centroid with the right index finger (Figure 1B). Centroid locations in the estimation task were distributed uniformly across the central 25% of the display area. We manipulated the type of feedback delivered during the training phase: observers in the vector condition (N = 15) received vector error feedback on a trial-by-trial basis (i.e., about the distance and direction of the centroid from their response), in addition to receiving scalar error feedback at the end of each block (i.e., proportional to the average deviation, without direction information); observers in the control condition (N = 15) received only the blockwise scalar error feedback. Vector error feedback was delivered by marking the correct centroid location with a green crosshair at the time of response. Scalar error feedback was delivered as point totals, where the number of points earned followed a Gaussian reward function of deviation from the centroid.

RESULTS AND DISCUSSION

Pretest phase performance on the discrimination task in both vector and control groups was significantly better than chance (58.4% correct vs. 25%), $t(29) = 18.4$, $p < .001$, and did not differ between groups, $t(28) = 1.32$, $p = .199$. Estimates of c, which controlled discrimination task difficulty, did not differ between groups, $t(28) = 1.08$, $p = .291$.

Training phase performance—measured as root mean squared error (RMSE)—was marginally better in the vector group than in the control group, $t(28) = 1.91$, $p = .067$ (Figure 1D). When we examined errors in the first of 12 training blocks (50 trials), this difference was more robust, $t(28) = 2.45$, $p = .021$. Verifying that the early advantage for the vector group was due to improvement across the first block rather than a baseline difference, a 2 (time bin: first vs. second half) × 2 (condition: vector vs. control) mixed measures ANOVA revealed a significant interaction, $F(1, 28) = 5.46$, $p = .027$. Importantly, performance did not initially differ between groups in the first half, $t(28) = 1.26$, $p = .219$. Taken together, these results are consistent with the possibility that that vector error feedback induces rapid calibration of the centroid representation, although further task practice does not necessarily reduce noise in centroid estimates—at least over the training period tested.

Posttest phase performance revealed the primary consequences of training: enhancement of perceptual discrimination was greater in the vector group (+7.1% correct), $t(14) = 3.28$, $p = .006$, than in the control group (+0.01% correct), $t(14) = 0.05$, $p > .900$, after controlling for pretest performance (ANCOVA), $F(1, 26) = 5.09$, $p = .033$ (Figure 1E). Thus, although vector error feedback in the training phase only marginally improved concurrent accuracy in the estimation task, this feedback appears to have tuned centroid representations in a more general manner, promoting transfer to a separate perceptual task that required a categorical decision.

In sum, our preliminary findings suggest that statistical summary representations are not merely an automatic consequence of visual experience. Rather, they can be rapidly tuned based on external feedback to flexibly support our perceptual goals in different contexts.

REFERENCES

Alvarez, G. A. (2011). Representing multiple objects as an ensemble enhances visual cognition. *Trends in Cognitive Sciences*, *15*(3), 122–131. doi:10.1016/j.tics.2011.01.003

Alvarez, G. A., & Oliva, A. (2008). The representation of simple ensemble visual features outside the focus of attention. *Psychological Science*, *19*(4), 392–398. doi:10.1111/j.1467-9280.2008.02098.x

Bauer, B. (2009). The danger of trial-by-trial knowledge of results in perceptual averaging studies. *Attention, Perception, and Psychophysics*, *71*(3), 655–665. doi:10.3758/APP.71.3.655

Chong, S. C., & Treisman, A. (2003). Representation of statistical properties. *Vision Research*, *43*(4), 393–404. doi:10.1016/S0042-6989(02)00596-5

Haberman, J., & Whitney, D. (2007). Rapid extraction of mean emotion and gender from sets of faces. *Current Biology*, *17*(17), R751–R753. doi:10.1016/j.cub.2007.06.039

Myczek, K., & Simons, D. J. (2008). Better than average: Alternatives to statistical summary representations for rapid judgments of average size. *Perception and Psychophysics*, *70*(5), 772–788. doi:10.3758/PP.70.5.772

Parkes, L., Lund, J., Angelucci, A., Solomon, J. A., & Morgan, M. (2001). Compulsory averaging of crowded orientation signals in human vision. *Nature Neuroscience*, *4*(7), 739–744. doi:10.1038/89532

Simoncelli, E. P., & Olshausen, B. A. (2001). Natural image statistics and neural representation. *Annual Review of Neuroscience*, *24*(1), 1193–1216. doi:10.1146/annurev.neuro.24.1.1193

Watson, A. B., & Pelli, D. G. (1983). QUEST: A Bayesian adaptive psychometric method. *Perception and Psychophysics*, *33*(2), 113–120. doi:10.3758/BF03202828