Perceptual learning of global pattern motion occurs on the basis of local motion

Shigeaki Nishina

Honda Research Institute Japan, Saitama, Japan, &
Department of Psychology, Boston University,
Boston, MA, USA

Mitsuo Kawato

ATR Computational Neuroscience Laboratories,
Kyoto, Japan

Takeo Watanabe

Department of Psychology, Boston University,
Boston, MA, USA

Enhancement in perceptual learning of a visual stimulus can often be explained either by learning of integrated visual information that is processed in higher visual areas or by learning of component information that is processed in lower visual areas. It is not clear on which visual information perceptual learning is predominantly based. We examined whether perceptual learning of global pattern motion occurs on the basis of local or global motion as a result of performance improvement in detecting contraction (or expansion) in a display in which contracting (or expanding) dots slightly outnumbers expanding (or contracting) dots. We measured the degree of transfer of the learning effect by presenting test stimuli spatially shifted so that the region of the test stimuli partially overlapped the trained region. The results showed that the degree of transfer was entirely dependent on how similar local motion directions in the test stimuli are to those in the trained stimulus in the overlapping area, irrespective of whether a test stimulus contained the same global motion direction as the trained or not. These results indicate that perceptual learning at least in the present setting occurs on the basis of local motion signals.

Keywords: perceptual learning, motion perception, global motion, local motion, MT, MST

Introduction

Perceptual learning is defined as improvement in performance with extensive training or exposure to a sensory feature (Fahle & Poggio, 2002). A number of studies have shown that perceptual learning occurs for visual motion (Ball & Sekuler, 1982, 1987; Law & Gold, 2008; Liu, 1999; Seitz & Watanabe, 2003; Watanabe, Náñez, & Sasaki, 2001). Although perceptual learning of motion has been broadly reported, it is still unclear on which stage of information processing perceptual learning of motion is based. Different types of motion best activate different brain areas with different response properties. Neurons in the primary visual cortex, for example, have smaller receptive fields and are highly responsive particularly to local motion signals. On the other hand, neurons in the higher areas such as monkey MT/MST or human MT+ have larger receptive fields. Especially, neurons in MST are known to respond best to more global motion (Dukelow et al., 2001; Komatsu & Wurtz, 1988a, 1988b; Tanaka et al., 1986).

In the present study, we address the question concerning whether a local or global motion processing stage is predominantly involved in improvement in performance of detecting global motion.

Experiment

Cells in MST of monkeys (Tanaka & Saito, 1989) and corresponding MT+ in humans respond well to global pattern motion (Dukelow et al., 2001; Koyama et al., 2005; Morrone et al., 2000). Cells in MST of monkeys selectively respond to either contraction or expansion motion (Tanaka & Saito, 1989). They have significantly larger receptive fields than those in V1 or MT. In addition, MST cells' response to contraction/expansion is largely location-invariant (Duffy & Wurtz, 1991) although a certain degree of location preference has been reported (Duffy & Wurtz, 1995).

A hypothesis is that the perceptual learning of global motion occurs on the basis of the global motion information that is processed in such brain areas as monkey MST or human MT+. If it is the case, the effect of learning will show strong global motion direction specificity and very weak location specificity. In other

words, we will observe strong transfer of the learning effect from a trained location to another location, if the same global motion direction is presented for training and test sessions. It is also predicted that little or no improvement will be observed when different global motion directions are used between training and test sessions.

On the other hand, if the perceptual learning occurs on the basis of local motion information that is processed where receptive fields are generally much smaller, strong location specificity and weak transfer of the learning effect over spatial locations should be observed.

This experiment was designed to address the question concerning whether the learning of contraction/expansion occurs on the basis of a global motion processing or not, by presenting the global motion at different two spatial locations for test and training stages. Those locations were partially overlapped and thus interactions between the trained and tested locations are systematically examined.

Subjects

Eleven subjects participated in this experiment. They were all naïve as to the purpose of the experiment and received payment for their completion of the experiment.

All had normal or corrected-to-normal vision. The experiment was conducted under the protocol approved by the internal review board of Boston University.

Apparatus and stimuli

The stimuli were presented using Psychophysics Toolbox (Brainard, 1997; Pelli, 1997) for MATLAB (The MathWorks, Natick, MA) running on a Macintosh G4 computer. The stimuli appeared on a Radius 21′ CRT monitor (1024 by 768 pixels, 100 HZ refresh rate) connected to the computer. The viewing distance was 0.66 m and one pixel size subtended 1.88 arcmin. A chin rest was used to maintain the subject's head position. The subjects used a numeric keypad to make a response. A fixation point was presented at the center of the screen. Dynamic random dots were presented in a circular area with 25 deg diameters, which was centered at 4 deg left or right from the fixation point (Figure 1A). The width and height of each dot was 5.64 arcmin and the density of the dots was 15 dots/deg^2. Each dot moved in either the inward (contraction) or outward (expansion) direction. In this study, we use the term "speed" as a signed value representing the motion along the radial direction. The
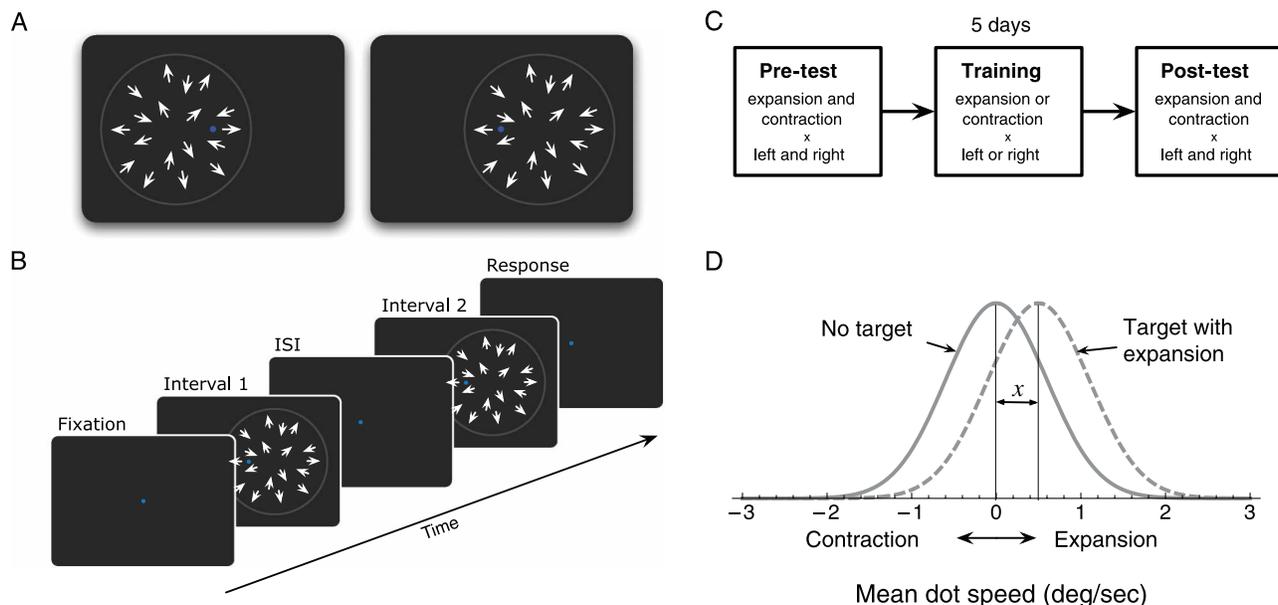

Figure 1. (A) Radially moving random dot pattern was presented as stimuli. In these schematic figures, the arrows show the directions of moving dots. Dots were presented in a circular area 4 deg shifted either to the left or to the right from the center. A blue fixation dot was presented at the center of the screen. The gray circles indicate the areas in which dots were presented, and were not shown in the actual experiments. (B) Stimulus sequence in a trial. (C) Three stages of the experiment. Each subject was trained only for one of the combinations of the two locations (left and right) and the two directions (inward and outward) throughout the training stage. All the four conditions (left/right × inward/outward) were tested in the test stages. (D) The motion direction and speed of each dot were randomly chosen based on a normal distribution as shown here. For the 2IFC task, for a target interval with expansion, the speed and direction (negative for contraction and positive for expansion) of a dot were randomly chosen from the distribution represented as the dotted curve with the mean speed at x. For no target intervals, the speed and direction of a dot were from the distribution represented as the solid curve with the mean speed at 0.

speed of each dot was assigned based on a normal distribution with 2 deg/sec standard deviation (Figure 1D, one-dimensional Gaussian along the radial direction), where negative and positive speeds correspond to contraction and expansion motion directions, respectively. Each dot had a lifetime of 20 frames (200 ms).

Procedure

The experiment consisted of three stages, pre-test, training, and post-test (Figure 1C). The task was basically the same for all the three stages. In each trial, two motion patterns were presented in two temporal intervals (Figure 1B). The duration of each interval was 300 ms. A blank screen was presented between the two intervals with 300 ms duration. The average dot speed was zero in the no target interval and slightly negative (contraction) or positive (expansion) in the target interval (Figure 1D). The participants were randomly divided into four subject groups, the expansion-trained at left, expansion-trained at right, contraction-trained at left, and contraction trained at right. During the 5-day training stage, the motion stimuli were consistently presented at the pre-assigned location. A direction of motion was also pre-assigned for each subject group and consistently presented in the target interval of each trial in the training stage. The subjects were instructed to indicate which of the two intervals contained the radial motion with shifted (non-zero) average by pressing one of two keys corresponding to the two intervals. Thus, the subjects could not respond on the basis of an individual dot motion. Thresholds of the mean dot speed for the target stimuli (x in Figure 1D) were measured using a staircase method. Each subject performed pre- and post-tests on separate days before and after the training stage, where thresholds for the four conditions (trained and untrained motion directions at left and right locations) were measured. Each session lasted approximately one hour.

Results

Figure 2 indicates the mean improvements (threshold in the pre-test − threshold in the post-test) for the four test conditions. The error bars represent the standard error of the mean across subjects. The largest improvement was shown in the trained condition in which the motion direction and the stimulus location were identical to those in the trained motion pattern (leftmost bar). The second largest improvement was observed for untrained motion + different location (rightmost bar). That is the condition in which the opposite global motion direction to the trained was presented on the opposite side to the trained motion pattern. If learning occurs on the basis of global motion,
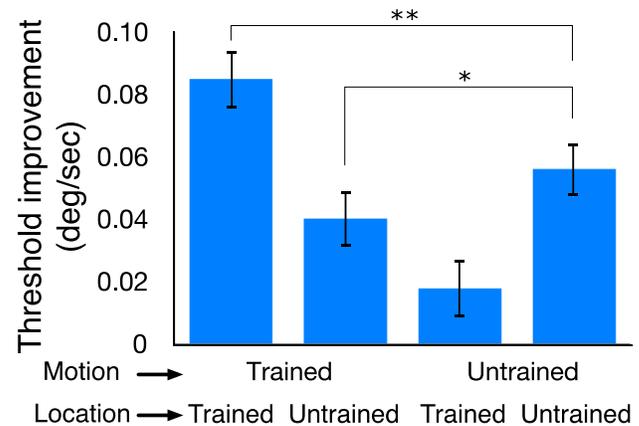

Figure 2. Performance improvements as threshold difference (post-test − pre-test) are shown for the four conditions. The error bars represent the standard error of the mean across subjects. The single and double asterisks indicate significance levels for the pairs, $p < 0.05$ and $p < 0.01$, respectively. Note that the significance levels are indicated only for the pairs mentioned in the text.

the improvement should be minimum for that condition because the tested global motion direction was opposite to the trained direction. However, significantly larger improvement was observed in untrained motion + different location (rightmost bar) than in trained motion + different location (second from the left) in which the motion direction was identical but locations were on the opposite side to the trained ($t(10) = 2.33$, $p < 0.05$, paired t-test performed on the whole subject group).

These results indicate that the improvement in a detection of global motion is not due to learning of the global motion but to learning of local dot motion while the subjects seem to have utilized the global motion presented in the larger area to accomplish the task.

Although MST neurons are known to have very large receptive field and their response are largely location invariant, some location preferred responses in monkey MSTd cells have also been reported (Duffy & Wurtz, 1995). However, this locational preference does not entirely explain our results. While the responses of the cells that show locational preference do predict the transfer to trained motion + different location, they do not predict the transfer to untrained motion + different location. In addition, the location specificity of these cells is not very strong (Duffy & Wurtz, 1995). Thus, our results can not be completely attributed to the location preference in MSTd cells.

The task could not be performed correctly if the subject paid attention to individual dot motion since the presence/absence of a global motion is statistically determined by a collection of local motion signals. In addition, if the subjects just attended to a small portion around the fixation and the learning occurred only within that area,

then the transfer to the untrained motion + different location should be almost perfect because the local motion shown in such area was almost identical. However the observed transfer to untrained motion + different location (rightmost in Figure 2) was significantly smaller ($t(10) = 3.22$, $p < 0.01$) than the trained motion + same location condition (leftmost). These results indicate that the subjects utilized motion information from a larger area. Furthermore, while the improvement for untrained motion + different location alone could be explained by assuming that the subjects learned to attend to peripheral locations, smaller improvement for trained motion + different location than for untrained motion + different location cannot. Also, the entire pattern of the results can not be explained by considering both locational preference of MSTd neurons and spatially focused attention, because transfer to an untrained global motion direction cannot be

predicted by learning on the basis of global pattern motion in any case.

Thus, we conclude that the results are at odds with the global motion hypothesis, but in accordance with the local motion hypothesis. In the next section, we propose a model in order to further understand the interactions between the trained and tested conditions.

Model and simulation

A possible explanation of the results is that in spite of the fact that the highest performance is expected when attending to a global motion, learning occurs on the basis of local motion signals. Namely, the presentation of a stimulus in which the number of one direction signals is
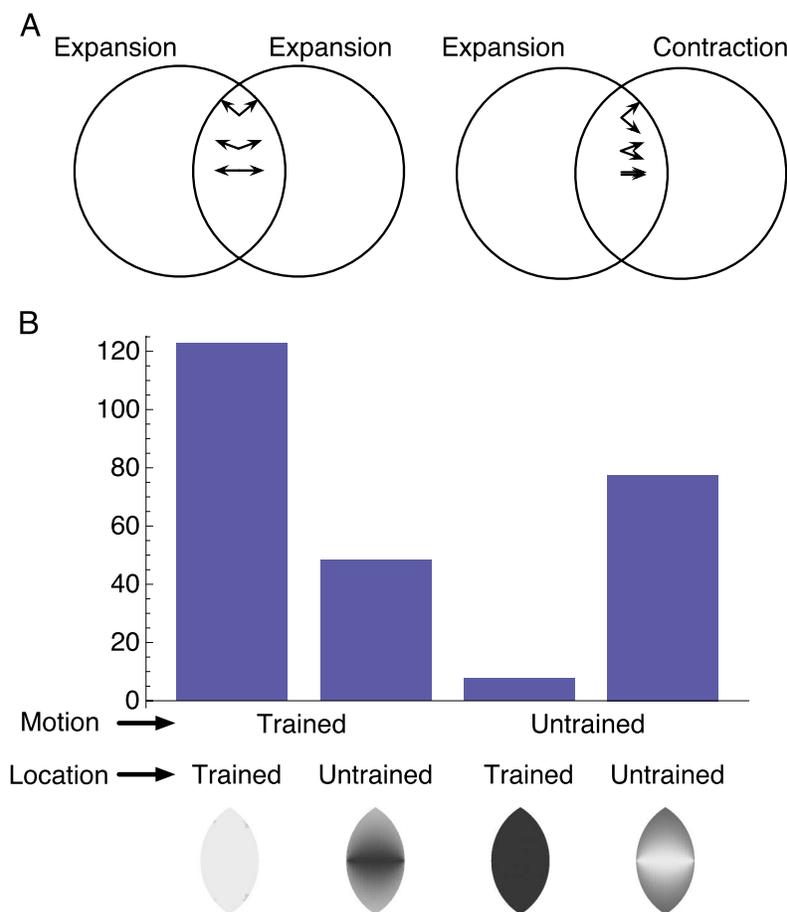

Figure 3. (A) Two possible interactions of local motions between a trained stimulus and a tested stimulus that are shifted but partially overlapped. In the left figure, expansion was trained on the left and tested on the right. In the right figure, expansion was trained on the left and contraction was tested on the right. The mean directional difference between local motion vectors in the overlapping area of the trained and tested stimuli is much larger in the former case than in the latter case. Thus, more facilitatory interactions are predicted in the latter case. (B) Predicted amounts of improvements for the four conditions in the experiment when In L was trained were calculated as local interactions between the trained and tested motion patterns. The bottom figures show the local interaction between the trained and tested stimuli mapped in the overlapped area. The gray scale corresponds to the amount of interaction from none (black) to the maximum (white).

larger than that of the opposite direction statistically enhances sensitivity of local motion detectors to the former direction than that to the opposite motion direction.

Assuming that the learning occurs in a local motion processing stage where each motion detector has a relatively small receptive field, the amount of the learning effect would be determined by how much trained local motion detectors are excited in the post-test stage compared to the pre-test stage. It can be assumed that a trained local motion detector is more excited when the direction of local motion presented in the receptive field of the trained detector is closer to the direction to which the detector is tuned. How well does this model predict the experimental results?

Figure 3A schematically shows directional differences of local motion vectors between a trained and tested stimuli when the tested and trained global locations were different. The arrows show directions of local motion at a few selected points. Local motion directions for the expansion motion (L) are represented as:

$$L_{\text{ExpL}}(x, y) = \tan^{-1}\left(x + \frac{r}{2}, y\right)$$
$$L_{\text{ExpR}}(x, y) = \tan^{-1}\left(x - \frac{r}{2}, y\right), \quad (1)$$

where (x, y) is the location of each point and r is the radius of the circular area in which the global motion stimulus was presented. Directions are opposite for the contraction motion:

$$L_{\text{ContL}}(x, y) = \tan^{-1}\left(x + \frac{r}{2}, y\right) + \pi$$
$$L_{\text{ContR}}(x, y) = \tan^{-1}\left(x - \frac{r}{2}, y\right) + \pi. \quad (2)$$

In Figure 3A, the local motion directions are closer between test and training when global directions are different (right), than when the global directions are the same (left). The interaction of the detectors tuned to two different directions can be defined by assuming a tuning function that has non-zero extent around the tuned direction. Here we assume that the extent is represented as a Gaussian function with 90 deg FWHM (full width at half maximum), which is a typical bandwidth shown neurophysiologically for MT neurons (Albright, 1984). Thus, the amount of interaction C at point (x, y) when contraction is trained on the left and contraction is tested on the right for example can be defined as:

$$C_{\text{ContL,ContR}}(x, y)$$
$$= \begin{cases} \dfrac{f(|L_{\text{ContL}}(x, y) - L_{\text{ContR}}(x, y)|)}{f(0)} & \left(x + \frac{r}{2}\right)^2 + y^2 < r^2 \wedge \left(x - \frac{r}{2}\right)^2 + y^2 < r^2 \\ 0 & \text{otherwise,} \end{cases}$$
$$(3)$$

where f is the Gaussian function. The total amount of interaction C_{TOTAL} is calculated by summing up interactions of local detectors at each location.

$$C_{\text{TOTAL,ContL,ContR}} = \sum_x \sum_y C_{\text{ContL,ContR}}(x, y). \quad (4)$$

Figure 3B shows the predicted total interactions for four test conditions. The results show a remarkable similarity to the results of the psychophysical experiment shown in Figure 2. The correlation between the psychophysical results and the simulated values was 0.949.

Discussion

In the present study, we examined whether improvement of detectability of a global motion (contraction or expansion) is based on learning of global motion or local motion. We found that the learning effect is highly location specific and that the observed transfer between different locations and the simulated values on the hypothesis that learning occurs on the basis of local motion were extremely highly correlated.

Since a seminal study by (Ball & Sekuler, 1982), a number of studies have been conducted to clarify mechanisms of perceptual learning of motion. Watanabe et al. (2002) have indicated that perceptual learning of motion as *task-irrelevant* occurs on the basis of local motion (Watanabe et al., 2002). However, it is controversial which stage is predominantly involved in perceptual learning of motion as *task-relevant*.

Our data strongly indicate that perceptual learning of global pattern motion occurs on the basis of local motion processing. This suggests that the learning is caused as a result of the local motion processing stage. However, we can not entirely rule out the possibility that a higher brain area that reads-out information from the local motion processing unit changes the way to read-out the information (Law & Gold, 2008; Mollon & Danilova, 1996).

The present results suggest that perceptual learning of motion at least in the present setting is highly likely to be based on changes related to local motion rather than global motion, although the task was to detect global motion.

Corresponding author: Shigeaki Nishina.

Email: nishina@jp.honda-ri.com.
Address: Honda Research Institute Japan, 8-1 Honcho, Wako-shi, Saitama 351-0188, Japan.

References

Albright, T. (1984). Direction and orientation selectivity of neurons in visual area MT of the macaque. *Journal of Neurophysiology, 52,* 1106–1130. [PubMed]

Ball, K., & Sekuler, R. (1982). A specific and enduring improvement in visual motion discrimination. *Science, 218,* 697–698. [PubMed]

Ball, K., & Sekuler, R. (1987). Direction-specific improvement in motion discrimination. *Vision Research, 27,* 953–965. [PubMed]

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

Duffy, C., & Wurtz, R. (1991). Sensitivity of MST neurons to optic flow stimuli: I. A continuum of response selectivity to large-field stimuli. *Journal of Neurophysiology, 65,* 1329–1345. [PubMed]

Duffy, C., & Wurtz, R. (1995). Response of monkey MST neurons to optic flow stimuli with shifted centers of motion. *Journal of Neuroscience, 15,* 5192–5208. [PubMed]

Dukelow, S., DeSouza, J., Culham, J., van den Berg, A., Menon, R., & Vilis, T. (2001). Distinguishing subregions of the human MT+ complex using visual fields and pursuit eye movements. *Journal of Neurophysiology, 86,* 1991–2000. [PubMed] [Article]

Fahle, M., & Poggio, T. (2002). *Perceptual learning.* Cambridge, MA: MIT Press.

Komatsu, H., & Wurtz, R. (1988a). Relation of cortical areas MT and MST to pursuit eye movements: I. Localization and visual properties of neurons. *Journal of Neurophysiology, 60,* 580–603. [PubMed]

Komatsu, H., & Wurtz, R. (1988b). Relation of cortical areas MT and MST to pursuit eye movements: III. Interaction with full-field visual stimulation. *Journal of Neurophysiology, 60,* 621–644. [PubMed]

Koyama, S., Sasaki, Y., Andersen, G., Tootell, R., Matsuura, M., & Watanabe, T. (2005). Separate processing of different global-motion structures in visual cortex is revealed by FMRI. *Current Biology, 15,* 2027–2032. [PubMed]

Law, C., & Gold, J. (2008). Neural correlates of perceptual learning in a sensory-motor, but not a sensory, cortical area. *Nature Neuroscience, 11,* 505–513. [PubMed] [Article]

Liu, Z. (1999). Perceptual learning in motion discrimination that generalizes across motion directions. *Proceedings of the National Academy of Sciences of the United States of America, 96,* 14085–14087. [PubMed] [Article]

Mollon, J., & Danilova, M. (1996). Three remarks on perceptual learning. *Spatial Vision, 10,* 51–58. [PubMed]

Morrone, M., Tosetti, M., Montanaro, D., Fiorentini, A., Cioni, G., & Burr, D. (2000). A cortical area that responds specifically to optic flow, revealed by fMRI. *Nature Neuroscience, 3,* 1322–1328. [PubMed]

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

Seitz, A., & Watanabe, T. (2003). Psychophysics: Is subliminal learning really passive? *Nature, 422,* 36. [PubMed]

Tanaka, K., Hikosaka, K., Saito, H., Yukie, M., Fukada, Y., & Iwai, E. (1986). Analysis of local and wide-field movements in the superior temporal visual areas of the macaque monkey. *Journal of Neuroscience, 6,* 134–144. [PubMed] [Article]

Tanaka, K., & Saito, H. (1989). Analysis of motion of the visual field by direction, expansion/contraction, and rotation cells clustered in the dorsal part of the medial superior temporal area of the macaque monkey. *Journal of Neurophysiology, 62,* 626–641. [PubMed]

Watanabe, T., Náñez, J. E., & Sasaki, Y. (2001). Perceptual learning without perception. *Nature, 413,* 844–848. [PubMed]

Watanabe, T., Náñez, J. E., Koyama, S., Mukai, I., Liederman, J., & Sasaki, Y. (2002). Greater plasticity in lower-level than higher-level visual motion processing in a passive perceptual learning task. *Nature Neuroscience, 5,* 1003–1009. [PubMed]