Interactions of Prefrontal Cortex in Relation to Awareness in Sensory Learning

Anthony Randal McIntosh,* M. Natasha Rajah, Nancy J. Lobaugh

In an associative learning paradigm, human subjects could be divided based on whether they were aware that one tone predicted a visual event and another did not. Only aware subjects acquired a differential behavioral response to the tones. Regional cerebral blood flow in left prefrontal cortex showed learning-related changes only in aware subjects. Left prefrontal cortex also showed changes in functional connectivity with contralateral prefrontal cortex, sensory association cortices, and cerebellum. Several of the interacting areas correlated with aware subjects' behavior. These results suggest cerebral processes underlying awareness are mediated through interactions of large-scale neurocognitive systems.

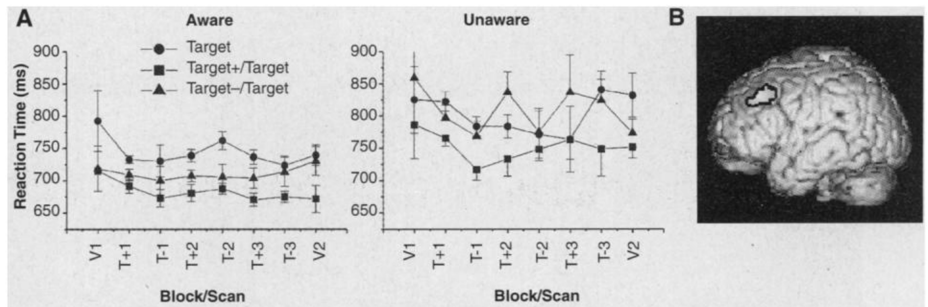
Many aspects of human behavior reflect influences of simple associative learning. Learning that a stimulus predicts the upcoming presence (or absence) of another event frequently occurs without overt awareness (*1*). In some cases, however, awareness of stimulus contingencies facilitates learning (*2*); in other cases, awareness appears to be necessary for learning (*3*). Learning with awareness typically recruits different brain regions than does learning without awareness (*4*), and prefrontal cortex is frequently implicated as a mediator of the awareness of associations (*5, 6*). Although integrity of key brain regions may be critical for awareness, some theories suggest awareness emerges from interactions within large-scale neural systems (*7, 8*).

We obtained evidence for neural system interactions related to awareness and performance in a positron emission tomography (PET) regional cerebral blood flow (rCBF) study of sensory associative learning. In the experiment, two tones and two visual stimuli were presented, and subjects responded to one visual stimulus (Target) and not to the other. One tone (Tone+) predicted an upcoming visual stimulus and the other tone (Tone–) predicted the absence of a visual stimulus (*9*). Based on post-experiment debriefing, subjects were classified as belonging to Group AWARE if they observed an association between the stimuli or used the tones to guide their behavior. Subjects who stated they did not notice any association between stimuli were classified as Group UNAWARE. Measures of rCBF taken in response to the Tone+ and Tone– three times across the experiment mapped the learning-related changes elicited by the tones (*10*).

Only AWARE subjects acquired differential reaction times (RT) on paired tone/target trials [repeated-measures analysis of variance (ANOVA); stimulus type by block interaction; $P < 0.001$; the group main effect was not significant (Fig. 1A)]. Initially, AWARE subjects showed facilitation to both tones (stimulus generalization), and behavioral separation between Tone+/Target and Tone–/Target trials emerged after block Tone+ 3. UNAWARE subjects showed no systematic RT differences across the experiment. Stimulus salience ratings corroborated the debriefing results. Both groups initially rated the Tone+ similarly. AWARE subjects rated the Tone+ as progressively more salient, and UNAWARE subjects rated it less salient (group by block interaction, $P < 0.05$). The Tone+ ratings between AWARE and UNAWARE subjects separated after Tone– 2, just before separation of RT. The temporal correspondence between RT differences and changes in salience measures suggests that AWARE subjects became aware of the stimulus contingencies between blocks Tone– 2 and Tone+ 3.

The strongest group difference in brain activity elicited by the tones was in left prefrontal cortex (LPFC) near Brodmann area 9 (*11*). In AWARE subjects, LPFC activity showed progressively greater activity to Tone– than Tone+ (Fig. 1C). Ventral and medial occipital cortices ($x = 22$, $y = -86$, $z = -20$; $x = -6$, $y = -88$, $z = 0$) and right thalamus ($x = 16$, $y = -22$, $z = 16$) showed progressively greater activity to Tone+ than to Tone– (*11*). In UNAWARE subjects, no consistent changes were seen in LPFC (Fig. 1C) or in any of the other regions. Our results support a prominent role for PFC in monitoring functions (*12*) and especially its putative role in awareness (*6, 13*). However, PFC activation has also been found in tasks where there was no overt awareness (*14*). It is therefore possible that interactions of PFC with other brain regions, present in AWARE but not in UNAWARE subjects, would better describe the neural system underlying awareness in this task.

To ascertain whether activity changes in LPFC reflected part of a large-scale neural system related to learning and awareness, we assessed the covariances, or functional connections (*15*), of LPFC. Scan-related changes in the covariances of the LPFC voxel with the



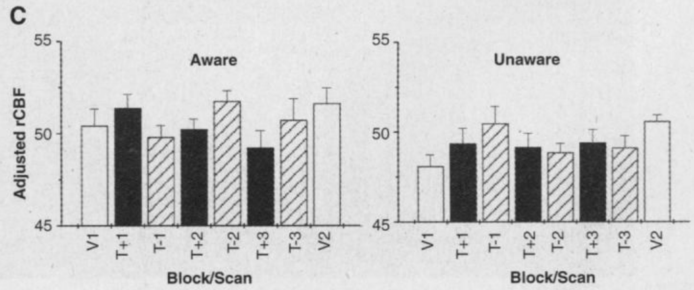


Fig. 1. (**A**) Average median reaction times preceding scans (±SE) for each group to trials of Target alone, Tone+/Target, and Tone–/Target. (**B**) Location and spatial extent of the LPFC region that showed significant group by task interactions ($x = -40$, $y = 20$, $z = 36$), overlaid on a structural magnetic resonance image. (**C**) Bar plots show average LPFC-adjusted rCBF (±SE) across training in the two groups. V, visual distractor scan; T+, Tone+ scan; T–, Tone– scan. Numbers associated with each block/scan indicate scan number for the particular stimulus.

Rotman Research Institute of Baycrest Centre, 3560 Bathurst Street, Toronto, Ontario M6A 2E1, Canada.

*To whom correspondence should be addressed. E-mail: mcintosh@psych.utoronto.ca

rest of the brain were tested using multivariate partial least squares (PLS) (*16*). If the learning-related changes in LPFC activity reflected systematic changes in interactions with other brain areas, then the pattern of functional connections should change across scans in AWARE subjects, but not in UNAWARE subjects.

In AWARE subjects, PLS analysis identified a set of regions whose functional connections with LPFC changed at the point where behavior began to differentiate (*17*). The dominant areas contributing to this pattern included right PFC, bilateral superior temporal cortices (auditory association), occipital cortex, and medial cerebellum (Table 1). Scan-specific correlation matrices of these areas and LPFC are presented in the left panels of Fig. 2. Correlations with LPFC strengthened on the Tone+ 2 scan, and the regions became highly intercorrelated. By Tone+ 3, when the discrimination was well learned, the interregional correlations were essentially unchanged and LPFC correlations were reversed. These effects indicate that the regions formed a coherent network differentially interacting with LPFC as learning proceeded in AWARE subjects.

A PLS analysis of brain-behavior relations in AWARE subjects was also performed to determine the regions most allied to changes in behavior and the degree of overlap with areas related to LPFC. Such correspondence would argue for a relation between awareness and performance, mediated through common functional connections with LPFC. The brain-behavior PLS analysis showed remarkable regional overlap with the PLS analysis of LPFC, as several of the areas in Table 1 were reliable in both analyses (*18*). The correlations of these regions with behavior are shown in the rightmost columns ("B") in Fig. 2.

The findings for the UNAWARE subjects contrast sharply with those of the AWARE subjects (Fig. 2, right panels). The correlation matrices for the same areas for UNAWARE

Table 1. Areas where activity covaried with LPFC activity and behavioral performance. The bootstrap ratio is the ratio of the parameter estimate from the PLS analysis for that voxel over its estimated standard error. Area number corresponds to the numbering scheme in Fig. 2

Bootstrap ratio		*x*	*y*	*z*	Atlas designation	Area number
LPFC	Behavior					
–	0.57	−40	20	36	Middle frontal, BA9	1
2.31	−2.35	−12	−62	−28	Cerebellum	2
2.41	−1.28	14	−64	−12	Lingual gyrus, BA18	3
2.06	−0.98	−18	−66	−12	Lingual gyrus, BA18	4
2.38	−2.17	0	−96	8	Cuneus, BA18	5
2.39	−2.42	14	−56	32	Cingulate, BA31	6
2.53	−2.01	−12	−52	32	Cingulate, BA31	7
−2.48	1.82	48	−36	−16	Middle temporal, BA21	8
−2.72	1.43	30	46	0	Middle frontal, BA46/10	9
−2.55	3.18	−60	−10	8	Superior temporal, BA22	10
−3.02	4.00	62	−12	8	Superior temporal, BA22	11
−2.64	2.41	48	18	32	Middle frontal, BA9	12
−2.40	1.79	32	−56	24	Inferior parietal, BA40	13

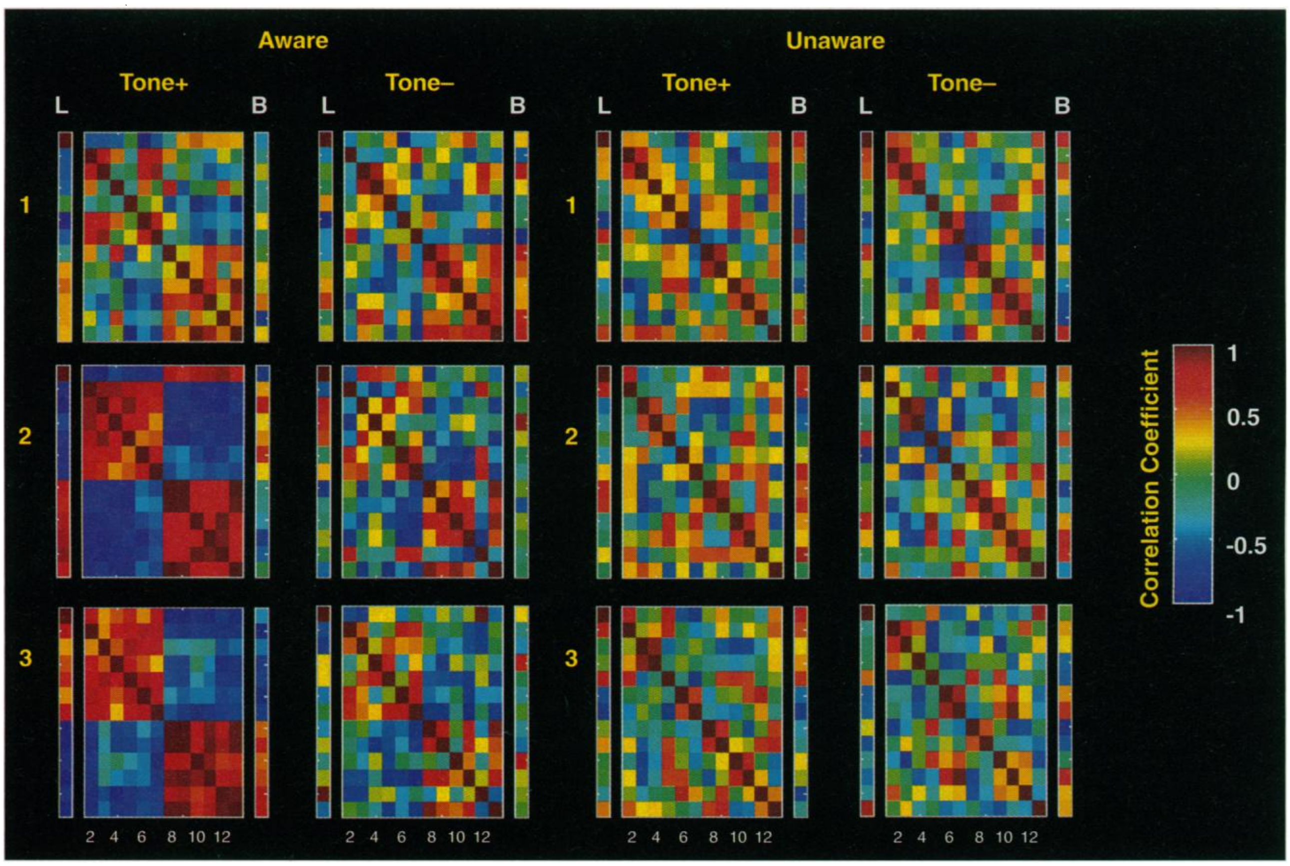


Fig. 2. Correlation matrices of LPFC with brain areas identified by PLS analysis in AWARE (left) and UNAWARE (right) subjects. The labels for each matrix correspond to scans, as in Fig. 1. The column numbers in each matrix correspond to the areas listed in Table 1. The correlation values are represented as color gradations (red = positive, blue = negative). The matrices are symmetric about the major diagonals, which are all 1's. The correlations of LPFC with the brain regions are shown in the first column (L) and the correlations of all regions with behavior are shown in the last column (B).

subjects showed minimal changes in the correlations with LPFC or behavior, and far less structure than for AWARE subjects.

We demonstrated that differences in awareness are related to learning and that both learning and awareness are related to LPFC activity. The behavioral effects were primarily expressed as a slowing of RT on Tone–/Target trials as the discrimination was learned. The functional connectivity profiles for the LPFC with the rest of the brain and for the brain with behavior demonstrated a change in covariance at the point in time where behavioral discrimination emerged and LPFC activity reversed. No such patterns were reliably observed for UNAWARE subjects.

The finding of high correlations among the regions related to LPFC is noteworthy. The increased coherence among these regions occurred simultaneously with a decrease in rCBF in LPFC. It emerges from these findings that interactions between LPFC and the system with which it interacted became more efficient as the discrimination was learned, resulting in a reduction of rCBF in response to the Tone+ (*19*). The switch between strong synchronous correlations on Tone+ scans to desynchrony during Tone– scans may reflect active suppression from LPFC, because Tone– predicted the absence of a visual stimulus (*20*). Other imaging studies noted regional activity decreases with learning, including PFC (*21*). Taken together, these data emphasize that learning involves both increases and decreases of brain activity.

These findings indicate that integration of specialized regions with larger neural systems is related to behavior and awareness. Specialization was demonstrated by differences in PFC activity between AWARE and UNAWARE subjects, consistent with a central role for PFC in monitoring sensory information and behavioral responses (*22*), as well as in processes whereby these operations enter into awareness. However, these areas did not act in isolation. Interactions of PFC and other specialized regions (for example, occipital and temporal regions) showed the strongest changes at the time when awareness likely emerged. Our findings indicate that awareness occurs through the integration of distributed regions. This does not preclude the possibility that PFC is critical for monitoring, but depending on how it interacts with other brain areas, these operations may or may not enter into awareness. In other words, PFC activation alone is not sufficient for awareness. Awareness, and in this case learning, may emerge only when this activity occurs together with specific interactions with other brain regions. Different patterns of PFC interactions led, in certain cases, to learning and awareness in AWARE subjects, and in other instances, to their absence in UNAWARE subjects. Such observations underscore the importance of examining neurocognitive organization in terms of changing interactions within large-scale networks (*8, 23*).

References and Notes

1. P. S. Wong, E. Bernat, S. Bunce, H. Shevrin, *Consciousness Cogn.* **6**, 519 (1997).
2. G. Davey, *Cognitive Processes and Pavlovian Conditioning in Humans* (Wiley, New York, 1987).
3. R. E. Clark and L. R. Squire, *Science* **280**, 77 (1998).
4. Greater activity in motor, dorsal prefrontal, and temporal cortex was observed in aware subjects in a sequence-learning task compared with unaware subjects, but learning and performance were equivalent in both groups [S. T. Grafton, E. Hazeltine, R. Ivry, *J. Cogn. Neurosci.* **7**, 497 (1995)].
5. D. T. Stuss, T. P. Picton, M. P. Alexander, in *The Frontal Lobes and Neuropsychiatric Illness*, S. Salloway, P. Malloy, J. Duffy, Eds. (American Psychiatric Press, Washington, DC, in press).
6. R. T. Knight, M. F. Grabowecky, D. Scabini, *Adv. Neurol.* **66**, 21 (1995).
7. F. Crick and C. Koch, *Nature* **375**, 121 (1995); T. W. Picton and D. T. Stuss, *Curr. Opin. Neurobiol.* **4**, 256 (1994); G. Tononi and G. M. Edelman, *Science* **282**, 1846 (1998).
8. M. M. Mesulam, *Brain* **121**, 1013 (1998).
9. Subjects (n = 12, seven males, 22 to 35 years, senior university students) pressed a keyboard button when a target stimulus appeared on a computer screen. The two visual stimuli were highly discriminable circles (13° visual angle) [A. R. McIntosh, R. Cabeza, N. J. Lobaugh, F. L. Bookstein, S. Houle, *Cereb. Cortex* **8**, 648 (1998)]. Subjects were informed that two tones would be presented through headphones (Tone+ = 2500 Hz FM, Tone– = 750 Hz FM, equated for perceived intensity), but were not instructed further about their significance. Stimulus durations were 500 ms; on paired auditory-visual trials, the auditory stimulus preceded the visual stimulus by 250 ms. The average intertrial interval was 7 s (range of 4 to 12 s). The eight 180-trial training blocks of ~11 min contained randomized sequences of paired auditory-visual trials mixed with unpaired presentations of all four stimuli. Scans immediately followed each training block and contained 10 presentations of one stimulus type. The first and last scans were of the visual distractor alone. The remaining six scans alternated between Tone+ alone and Tone– alone. Thus, because the visual target was never presented during a scan, motor responses were not required during any scans. The Tone+ predicted either visual stimulus on 77% of trials, and the Tone– predicted either visual stimulus on 11% of trials. These probabilities were used to ensure that the presentation of isolated stimuli during the scans was not a novel event. A 1-min break period followed each scan, and subjects rated the subjective salience of each stimulus. After the experiment, subjects were questioned to determine whether they were aware of stimulus contingencies and then debriefed. Six subjects were classified as AWARE and six as UNAWARE; there were no group differences in sex, age, education, or apparent motivation. Written informed consent was obtained, and subjects were paid. The Human Subjects Use Committee of Baycrest Centre approved the protocol.
10. Eight PET scans were conducted following bolus injections of 30 mCi [^{15}O]H_2O for each scan, using radioactive counts to approximate rCBF. Images were corrected for movement using AIR 3.0 [R. P. Woods, S. R. Cherry, J. C. Mazziotta, *J. Comput. Assisted Tomogr.* **16**, 620 (1992)], transformed to a PET rCBF template conforming to a standard human brain atlas [J. Talairach and P. Tournoux, *Co-Planar Stereotaxic Atlas of the Human Brain* (Thieme, New York, 1988)], and smoothed with a 10-mm isotropic Gaussian filter using SPM95 [K. J. Friston *et al.*, *Hum. Brain Mapp.* **2**, 165 (1995)].
11. The functional images were compared between groups using a repeated-measures ANOVA with permutation tests to assess significance at each voxel (uncorrected $P < 0.001$). The design matrix included contrasts for a general time effect across all eight scans, one contrasting the difference between auditory and visual stimulation, two acquisition contrasts comparing Tone+ and Tone– differences, and the interactions of these with group. The voxels that were significant for the group by acquisition interaction are discussed.
12. T. Shallice and P. Burgess, in *Frontal Lobe Function and Dysfunction*, H. S. Levin, H. M. Eisenberg, A. L. Benton, Eds. (Oxford Univ. Press, New York, 1991), pp. 125–138; D. T. Stuss and D. F. Benson, in *The Frontal Lobes Revisited*, E. Perecman, Ed. (Institute for Research on Behavioral Neuroscience Press, New York, 1987), pp. 141–158.
13. M. Petrides, *Neuropsychologia* **23**, 601 (1985).
14. G. S. Berns, J. D. Cohen, M. A. Mintun, *Science* **276**, 1272 (1997).
15. K. J. Friston, C. D. Frith, P. F. Liddle, R. S. J. Frackowiak, *J. Cereb. Blood Flow Metab.* **13**, 5 (1993).
16. A. R. McIntosh, F. L. Bookstein, J. V. Haxby, C. L. Grady, *Neuroimage* **3**, 143 (1996); A. R. McIntosh, L. Nyberg, F. Bookstein, E. Tulving, *Hum. Brain Mapp.* **5**, 323 (1997).
17. Covariances of the LPFC voxel with the rest of the image voxels were computed within each of the six tone scans, across subjects and between groups. Singular value decomposition of these covariance maps produces latent variables containing scan profiles and "singular" images. The scan profiles index changes in the covariances of LPFC with the regions identified in the singular image across scans and between groups. Brain–behavior relations were examined separately in the two groups. This analysis was identical to the analysis above, except that RT differences between Tone+/Target and Tone–/Target trials in each block were used to generate covariance images. Scan profiles index the changes in covariances of behavior with the singular image. The significance of the scan profiles within and between groups and the voxel contributions to the singular image was assessed using bootstrapping ($P < 0.01$) [A. R. Braun *et al.*, *Science* **279**, 91 (1998); A. R. McIntosh and F. Gonzalez-Lima, *J. Neurophysiol.* **80**, 3148 (1998)].
18. The correlation between the singular images from the LPFC–brain PLS and the brain–behavior PLS analyses was 0.68, and the 95% confidence interval extended from 0.10 to 0.85.
19. C. Büchel, J. T. Coull, and K. J. Friston [*Science* **283**, 1538 (1999)] noted an increase in neural interactions and decreased activity with learning among cortical visual areas.
20. Analysis of direct and indirect effects among the regions would be required to address this question {for example, effective connectivity [K. J. Friston, C. D. Frith, R. Fracowiak, *Hum. Brain Mapp.* **1**, 69 (1993)]}. Structural equation modeling [A. R. McIntosh *et al.*, *J. Neurosci.* **14**, 655 (1994)] of the connections of LPFC with auditory and visual cortices in AWARE subjects indicated strong influences of LPFC early in acquisition to the Tone+. By the end of training, effects on Tone– trials were stronger than on Tone+ trials. UNAWARE subjects showed no systematic relations. Sample size prohibits a full examination of these relations, so we regard them as tentative.
21. S. E. Molchan, T. Sunderland, A. R. McIntosh, P. Herscovitch, B. G. Schreurs, *Proc. Natl. Acad. Sci. U.S.A.* **91**, 8122 (1994); J. S. Morris, K. J. Friston, R. J. Dolan, *Proc. R. Soc. London Ser. B* **265**, 649 (1998); M. E. Raichle *et al.*, *Cereb. Cortex* **4**, 8 (1994); K. Sakai *et al.*, *J. Neurosci.* **18**, 1827 (1998).
22. M. Petrides, B. Alivisatos, A. C. Evans, E. Meyer, *Proc. Natl. Acad. Sci. U.S.A.* **90**, 873 (1993).
23. K. J. Friston, *Neuroimage* **5**, 213 (1997); A. R. McIntosh, *Ann. N.Y. Acad. Sci.* **855**, 556 (1998).

24. We thank the staff of the PET Centre at the Clarke Institute of Psychiatry, University of Toronto, for their valuable assistance, and E. Tulving, T. W. Picton, and M. J. Taylor for comments on the manuscript. Supported by Natural Sciences and Engineering Council of Canada grant OGP017034 and Medical Research Council of Canada grant MT-13623 to A.R.M.

3 February 1999; accepted 8 April 1999