Exp Brain Res (2005) 166: 170–183
DOI 10.1007/s00221-005-2351-5

RESEARCH ARTICLE

R. C. Miall · E. W. Jenkinson

Functional imaging of changes in cerebellar activity related to learning during a novel eye–hand tracking task

Received: 15 November 2004 / Accepted: 11 March 2005 / Published online: 5 August 2005

Abstract Coordination between the eyes and the hand is likely to be based on a process of motor learning, so that the interactions between the two systems can be accurately controlled. By using an unusual tracking task we measured the change in brain activation levels, as recorded with 3T functional magnetic resonance imaging (fMRI), between naïve human subjects and the same subjects after a period of extended training. Initially the performance of the two groups was similar. One subject group was then trained in a synchronous, coordinated, eye–hand task; the other group trained with a 304 ms temporal offset between hand and eye tracking movements. After training, different patterns of performance were observed for the groups, and different functional activation profiles. Significant change in the relationship between functional activation levels and eye–hand task conditions was predominantly restricted to visuo-motor areas of the lateral and vermal cerebellum. In an additional test with one of the subject groups, we show that there was increased cerebellar activation after learning, irrespective of change in performance error. These results suggest that two factors contribute to the measured blood oxygen level-dependent (BOLD) signal. One declined with training and may be directly related to performance error. The other increased after training, in the test conditions nearest to the training condition, and may therefore be related to acquisition of experience in the task. The loci of activity changes suggest that improved performance is because of selective modified processing of ocular and manual control signals within the cerebellum. These results support the suggestion that coordination between eye and hand movement is based on an internal model acquired by the cerebellum that provides predictive signals linking the control of the two effectors.

Keywords Human · Coordination · Functional imaging · Cerebellum · Motor learning

R. C. Miall
Behavioural Brain Sciences, School of Psychology,
University of Birmingham, Birmingham, B15 2TT, UK

R. C. Miall (✉) · E. W. Jenkinson
University Laboratory of Physiology, Parks Road,
Oxford, OX1 3PT, UK
E-mail: r.c.miall@bham.ac.uk
Tel.: +44-121-4142867
Fax: +44-121-4144897

Introduction

Many imaging groups have reported changes in functional activation of sensory-motor systems in the human brain that are contingent on motor learning or adaptation. A critical region in most of these studies is the cerebellum (Jenkins et al. 1994; Shadmehr and Holcomb 1997; Jueptner et al. 1997a, b; Tamada et al. 1999; Imamizu et al. 2000; Ramnani et al. 2000; Nezafat et al. 2001; Ramnani and Passingham 2001; Seidler et al. 2002; Imamizu et al. 2003), well recognised as important for motor learning (Thach et al. 1992; Ito 2002). Two patterns of change in activation are prominent. First, there is initial strong signal during early learning that decays as performance improves (Jueptner et al. 1997a; Imamizu et al. 2000; Nezafat et al. 2001). This is likely to be related to neural processing of motor errors: directly related to cerebellar signalling of error; to neuronal adaptation driven by the error; or to the increased motor output as subjects correct mistakes or co-contract to increase limb stiffness. Second, there is evidence for increased cerebellar activation which develops as learning proceeds, suggested to be a signature of an internal model or motor memory (Shadmehr and Holcomb 1997; Imamizu et al. 2000; Imamizu et al. 2003).

Functional activation of the cerebellum is, however, always combined with activity in other cerebral sensory-motor regions. Changes in activity caused by learning are, moreover, found both in cerebral and cerebellar regions. Because of this, when combined with the fact

that the cerebellum is sensitive to many different characteristics of task performance (e.g. speed, rate of movement, force, extent, timing, and accuracy), we cannot predict what components of the cerebellar responses are modified, nor whether changes in cerebellar function cause—or are caused by—activation changes in cerebral areas. Last, the areas in the cerebellum in which significant activation differences are observed during learning are not always straightforwardly mapped on to areas that are activated during the relevant hand or arm actions (e.g. Nezafat et al. 2001; Seidler et al. 2002; Imamizu et al. 2003).

We recently reported (Miall et al. 2001) that in a visually guided tracking task, the cerebellar blood oxygen level-dependent (BOLD) signal measured with functional magnetic resonance imaging (fMRI) was related to the temporal relationship between eye and hand movements. In fact, the quadratic relationship between eye–hand asynchrony and BOLD signal was found to be significant only in the cerebellum, even though performance of the task strongly activated the whole cerebrocerebellar circuit responsible for visually guided action. This task, manipulating the temporal relationships between eye and hand, might therefore provide a powerful tool to explore learning-dependent changes isolated to the cerebellum alone.

We report here three new experiments to test the relationship between tracking performance and BOLD signal in the cerebellum. Two groups were scanned before and after learning a specific eye–hand coordination condition. We first tested if learning-dependent changes in tracking performance are coupled with specific changes in the pattern of activity within the cerebellum or elsewhere. Second, we tested the predictive relationship between measures of performance and coordination and the BOLD signal. Third, we measured activity under constant performance conditions, to dissociate the hypothesized increase in cerebellar activation caused by learning from the decrease caused by reduction in error.

Methods

Subjects

Twenty-three subjects participated in these experiments after giving written informed consent, in accordance with Declaration of Helsinki. The Oxfordshire Clinical Research Ethics Committee approved the experiments. One subject failed to perform the tracking tasks adequately and is not considered further. One group (group A, $n = 10$: four male, six female; age range 19–21 years) was tested in experiment 1 only. Group B ($n = 12$: nine male, three female; age range 18–34, mean 23.3 years) were tested in experiment 2; and nine subjects from this group were also tested in experiment 3 (six male, three female, mean age 24.2).

Tracking and training paradigms

The details of the task have been published previously (Miall et al. 2001). In brief, subjects were required to perform ocular tracking of a visual target at the same time as performing manual compensatory tracking using a hand-held joystick. They used 90-degree prismatic

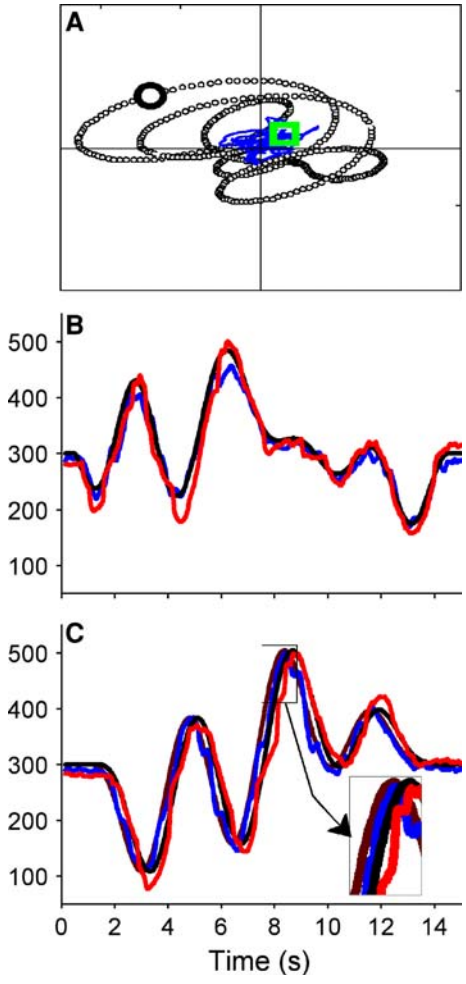

Fig. 1 Tracking task. **A**. Subjects viewed a display screen with a stationary cross-hair and two moving symbols (*circle* and *square*). The circle was the target for eye movement, and followed a smooth unpredictable trajectory (shown here, *black dots*, but not displayed on screen). The square was controlled by a hand-held joystick, and its distance from the cross-hairs reflected the error between desired and actual joystick positions. By compensating for the cursor's imposed displacement away from the centre following a similar target trajectory as for the eyes, subjects maintained the square cursor close to the cross hairs (blue trajectory, also not displayed on screen). **B**. Typical ocular and manual performance measured outside the fMRI scanner. The horizontal components of the target (*black*), eye (*red*) and joystick (*blue*) trajectories are plotted against time. The vertical axis is the horizontal screen coordinate (in pixels). In this condition, eye and hand target functions were identical and synchronous, and both eye and hand accurately follow the same spatial path. **C**. An example of the −304 ms offset condition in which the manual target (*brown*) anticipated the ocular target (*black*) by 304 ms. The joystick trajectory (*blue*) also precedes the ocular trajectory (*red*). The inset is a magnified segment

glasses to view a computer-generated display (Fig. 1A), projected on to a viewing screen at 640×480 resolution, with 70 Hz refresh rate. If necessary, subjects' vision while in the scanner was corrected with self-selected lenses. They held a lightweight joystick in their right hand and movement of the joystick in two dimensions was sampled at 26 Hz.

The tracking task for all experiments required ocular and manual responses. The target for eye movements was a white circle, 0.2° in diameter, moving in a smooth pseudo-random trajectory in two dimensions. The target position for the joystick-controlled cursor was given by a large centrally positioned stationary cross-hair (Fig. 1A), but the cursor was continually moving, following another pseudorandom trajectory, thus requiring so-called "compensatory" tracking responses with the joystick to bring the cursor back on to the stationary cross-hair. Movement of the tip of the joystick of approximately 6 cm (75° of joystick motion) was required, typically involving thumb, finger, and wrist movements. The target functions controlling the eye target and the cursor motion away from the cross-hair were randomised each trial (Miall et al. 2001). In the null or baseline condition ($\tau = 0$ ms) these were identical, inverted, versions of each other. Thus during ocular pursuit and accurate manual compensatory tracking, the subject's eye and hand would actually follow trajectories that were synchronous and consistent—as the eyes moved leftwards, for example, so should the hand, to compensate for the rightwards motion of the cursor away from the centre (Fig. 1B).

Temporal offsets between the eye and hand target functions were then introduced to parametrically vary the amount of eye–hand coordination across five test conditions ($\tau = -304$, -152, 0, $+152$, and $+304$ ms). Positive offsets ($\tau = +152$ or $+304$ ms) indicate the ocular target preceded the hand target; negative offsets ($\tau = -152$ or -304 ms) indicate the ocular target lagged behind the hand target (Fig. 1C). Hence, during accurate tracking, the eye and hand followed the same spatial path in respect of the display screen, but with a short time offset between the two trajectories. Experiments 1 and 3 also included an "independent" condition in which the target trajectories were spatially and temporally unrelated to each other (Miall et al. 2001).

Eye movement could not be recorded in the MR environment but was observed in all subjects during training in the laboratory. Eye movements were also recorded in other subjects, using an ASL 501 infrared eye tracker (see "Results"). Accurate ocular tracking required eye movement of up to 10° horizontally, 7.5° vertically, with maximum speeds of approximately 6 s^{-1}. Manual tracking was quantified by calculating the RMS error between the cursor and the central cross hairs, measured in screen pixels (1 pixel≈0.016 degrees at the eye). Because the target trajectories were randomised each 18 s trial, and thus had slightly different average speeds, RMS error varied from trial to trial. Hence the RMS errors were normalized by division of the mean RMS error per trial by mean target speed per trial (Miall et al. 2001) and are presented in arbitrary units.

All subjects were scanned with fMRI in two sessions 7 days apart, before and after a period of extended training that took place outside the MR scanner. One or two days before the first scanning session, all subjects were given limited training in the laboratory. This introduced them to the task, confirmed that they performed according to instruction, including maintaining their gaze on the ocular target as assessed by the observing experimenter, and provided approximately 3 min practice at the coordinated and independent conditions. This was followed by 12 min practice (40 trials) at the coordinated and $\tau = \pm 152$ ms temporal offset conditions (10 trials each plus rest, in pseudo-random order). A further 5–10 min practice of these last three conditions was provided in the scanner, while setting up for functional imaging. To keep subjects as naïve as possible we restricted practice to these three conditions, avoiding exposure to the more obvious and extreme conditions of ± 304 ms that were used in the experiments. Results show that learning in this task is slow, and we do not believe the limited practice of these conditions would have a major effect on performance.

After the first scanning session, subjects trained in the laboratory for 1 h a day for 4 days, lying in a supine position approximating their position in the scanner. Training consisted of three 20 min sessions each day, with a few minutes rest between sessions. No training was undertaken over the weekend; no performance differences were seen that related to the scheduling of practice and scanning sessions (Monday scans (with training on Tuesday–Friday) $n = 8$; Thursdays scans (training on Friday and Monday–Wednesday) $n = 6$; Friday scans (training on Monday–Thursday), $n = 9$).

Imaging

T2*-weighted echo-planar images were acquired for each subject using a 3T Siemens Vision scanner with a GEM BEST sequence. The field of view covered the whole brain: 256×256×125 mm, 64×64 voxels, 25 axial slices; TR = 3 s, TE = 30 ms, flip angle = 90°. For experiment 1, 427 volumes were acquired (total 21 min 21 s); 294 volumes were acquired for experiment 2 (14 min 24 s); and 216 volumes for experiment 3 (10 min 48 s). High-resolution T1-weighted structural images were also acquired for each subject.

Image analysis

General analysis procedures are described first (Smith 2001). Specifics for each experiment are given below. All analysis was carried out using FMRI Expert Analysis

Tool (FEAT) Version 5.00, part of FSL (''FMRIB's Software Library''; a full description of the functions and processes used in fMRI analysis are available at http://www.fmrib.ox.ac.uk/fsl). In brief, the times series data were motion-corrected and spatially and temporally smoothed (FWHM 5 mm; high pass cut-off 162 s). Statistical analysis was then performed using general linear modelling with FILM (''FMRIB's Improved Linear Model'') with local autocorrelation correction (Friston et al. 1994; Woolrich et al. 2001). All statistical parametric maps were combined across subjects with least squares mixed effects models (also called random effects analyses). *Z*-score statistical images (Gaussian T-statistics) were thresholded using clusters determined by $Z > 2.3$ and a corrected cluster significance threshold of $P = 0.01$ (Worsley et al. 1992; Friston et al. 1994; Forman et al. 1995). Registration of the functional activation maps to high resolution and standard structural images was carried out using FLIRT (''FMRIB's Linear Image Registration Tool'') (Jenkinson and Smith 2001; Jenkinson et al. 2002), using an affine 12 degrees of freedom transformation.

Significant activation clusters were rendered as colour images on to the median brain image of the group. Locations of cluster maxima and local maxima are reported; identification of maxima was examined in the median images by reference to published atlases (Talairach and Tournoux 1988; Schmahmann et al. 1999; Duvernoy 1999). Parameter estimates (PEs) from the general linear model for each subject were extracted from the data using Featquery, extracting the within-subject mean values across those voxels identified as regions of significant activation in the mixed effect analyses. Group statistics on these estimates were performed with SPSS (SPSS, Chicago, Illinois, USA).

Specific experimental details

We report three experiments. Subject group A performed experiment 1. Group B performed experiment 2, nine subjects tested in group B were also tested in experiment 3. The scanning for experiment 3 took place immediately after the scans for experiment 2, both before and after training.

Experiment 1: training with eye–hand offsets

Ten subjects (group A) performed six tracking conditions, presented 10 times each in blocks of 18 s. These were the five temporal offset conditions described above, and the independent eye–hand tracking condition (see ''Tracking and training paradigms'', above, and Miall et al. (2001); Miall and Reckess (2002)). A rest or fixation condition was presented every seventh block, when the circular and square icons turned red and remained stationary; the other six active tracking conditions were pseudorandomly ordered. Between the two scan sessions, subjects in group A were trained exclusively at the $\tau = -304$ ms condition (with the ocular target lagging behind the hand target).

The general linear models used to analyse experiment 1 consisted of six categorical variables defining the six active tracking conditions as explanatory variables (EVs). Temporal derivatives of these variables were also included as covariates to enable temporal adjustment of the EV fits to the data. Because randomisation of the target motion every 18 s trial introduced speed variations between trials, which would affect activation levels (Turner et al. 1998; Imamizu et al. 2000; Miall et al. 2001), the average speed of the joystick measured every 3 s, and its temporal derivative, were also used as covariates. All the above 14 covariates were convolved with a gamma function (SD 3 s, mean lag 6 s) to model the haemodynamic responses. In addition, six head motion correction parameters (x, y, z and rotations) were included as covariates, but without convolution by the gamma function, to model the direct position-dependent effects of head-motion on the recorded signal. Thus a total of 20 EVs were used in the multiple regression of the BOLD signal, producing an array of PEs, or weights, of which only the six are of interest. Voxels with PEs significantly greater than zero define regions where activation is greater than during the rest condition.

Regions of activation parametrically related to the temporal offset between the two target functions were then identified using a vector of six non-linear contrast weights (CW) of ($+5$, -5, $+1$, $+3$, $+1$, and -5), weighting the independent condition and temporal offsets of $\tau = -304$, -152, 0, $+152$, and $+304$ ms, respectively. This model tests for activation quadratically related to the temporal offsets, combined with high activation during independent tracking as predicted by our earlier work (Miall et al. 2001). Thus the contrast identified voxels for which the dot product was significantly greater than zero.

$$CW * PE > 0.0 \quad (1)$$

Comparisons of the main contrasts (the estimates of PE for the six tracking conditions) between sessions 1 and 2 and the contrast of the parametrically related activation between sessions (Eq. 3) were tested with paired *t*-tests using mixed effects analyses, which are also known as random effects analyses.

Experiment 2: training with coordinated eye–hand action

A separate group of 12 subjects (group B) performed five tracking conditions, presented eight times each in blocks of 18 s. A rest or fixation condition was presented every sixth block. The five active conditions were pseudorandomly ordered and consisted of the five temporal offset conditions (see ''Tracking paradigm'', above). Between

the two scan sessions, this subject group was trained at the null, or coordinated ($\tau = 0$ ms) condition.

The GLM analysis was basically identical to that used in experiment 1, except that 18 instead of 20 EVs were used, because there was one fewer tracking condition and one fewer temporal derivative. Regions of activation quadratically related to the temporal offset between the two target functions were identified using five non-linear contrast weights of CW = (−4, +2, + 4, +2, and −4). These weights differ from those for experiment 1 only because in the first experiment an extra first parameter was needed to code for the independent condition and then, to maintain a mean of zero, all weights were offset by −1. However, they test the same relative difference across the five temporal offset conditions. Comparisons between sessions were made in the same way as in experiment 1.

Experiment 3: constant error tracking

These scanning conditions immediately followed experiment 2 (before and after training), and were completed by nine of the 12 subjects from group B. Three conditions were presented 10 times each, in blocks of 18 s. Every seventh block was rest (fixation), the other three conditions were presented in pseudorandom order, twice in each set of six blocks. These were the coordinated ($\tau = 0$ ms), the asynchronous ($\tau = -304$ ms), and the independent eye–hand conditions (Experiment 1). However, an "error-clamping" algorithm was used in which subject performance was measured trial-by-trial, and the mean target speed for each 18 s trial was adjusted by the ratio of RMS tracking error recorded in the previous trial of the same condition versus a constant target performance level of 7.0 (arbitrary units; Fig. 3). This was designed to maintain constant levels of performance (RMS error) across tracking conditions and across pre-learning and post-learning sessions. For analysis, the first trial of each condition was discarded. The error clamping procedure failed to hold errors at the target level during the independent tracking condition (mean RMS error 10% higher than target; SD of within-subject RMS errors 62% greater than for the other two conditions). This condition was therefore discarded from all further analysis.

For experiment 3, comparisons between the main contrasts of sessions 1 and 2 (the two tracking conditions with respect to rest, the movement speed covariate, and the combined contrast for tracking-related activation) were made in the same way as in experiment 1. Because of the small group size ($n = 9$), these statistical maps were spatially restricted to the cerebellum before thresholding for significant clusters, using a mask based

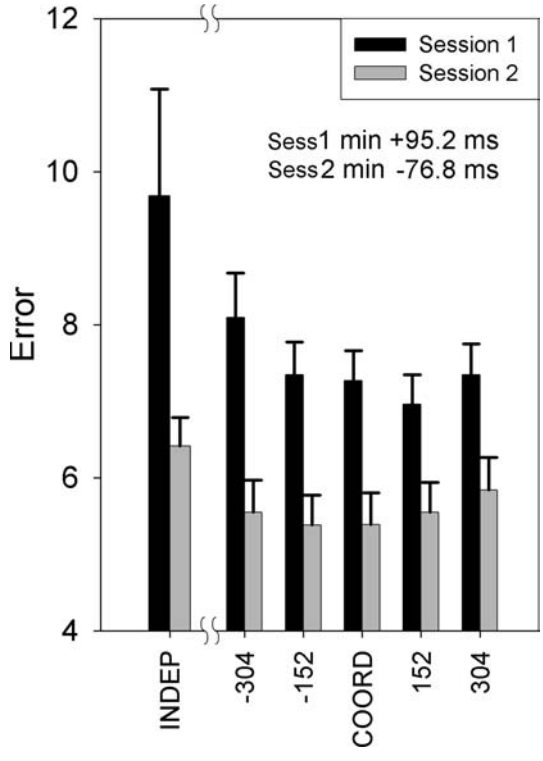

Fig. 2 Tracking performance, group A, experiment 1. Mean RMS tracking errors (*bars*, ±1 SEM, $n = 10$) were measured from the joystick position signal recorded throughout each scanning session, normalized for the mean target speed (see "Methods"). This group was tested in five eye–hand temporal offset conditions and in an independent eye–hand condition, and trained at the $\tau = -304$ ms asynchronous condition between sessions 1 and 2. Optimum eye–hand asynchrony was estimated for each session from quadratic regression across the five temporal offset conditions

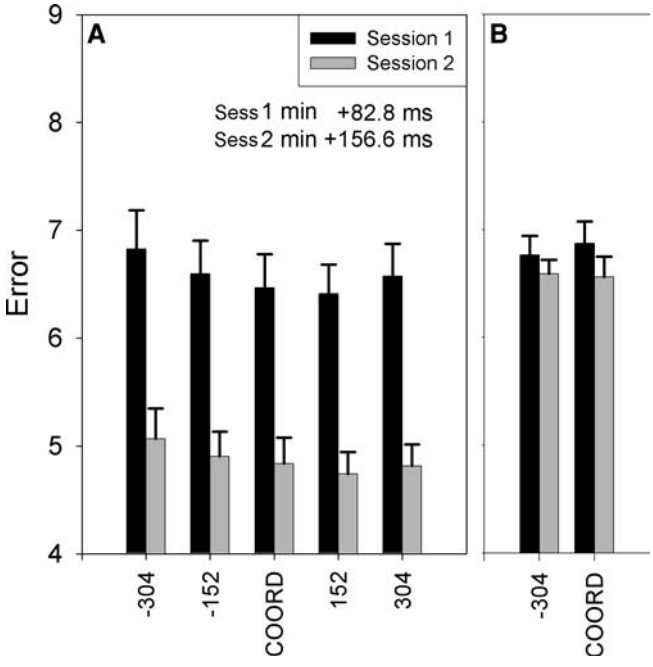

Fig. 3 Tracking performance, group B, experiments 2 and 3. This group was tested in five eye–hand temporal offset conditions, and trained at the coordinated, synchronous, condition. **A**. Mean RMS tracking errors from experiment 2 (±1 SEM, $n = 12$) are presented in the same format as for Fig. 2. **B**. Mean RMS tracking errors from experiment 3 (±1 SEM, $n = 9$), during on-line clamping of RMS error by changing target speed

on the median structural image. This restricted the number of voxels tested in the analysis, so increasing statistical power.

BOLD regression analysis

A multiple regression equation was used to estimate the predicted BOLD signal level, B, by fitting the group mean BOLD values calculated over cerebellar regions of interest (i.e. the areas of significant difference in Figs. 6A, B). The regression equation was:

$$B_{G,S,\tau} = \alpha E_{G,S,\tau} + \beta_S C^2_{\tau - k_{G,S}} + \gamma_S \quad (2)$$

where S is scanning session (1 or 2), G is subject group (A or B), τ is the temporal offset for a given condition. The term $k_{G,S}$ is the optimum temporal offset for that group and session; this is the estimated point at which the performance versus temporal offset curve reaches a minimum (Figs. 2 and 3A). The error term ($E_{G,S,\tau}$) was the mean performance error for a given group, in a specific session and offset condition; the "coordination" term ($C^2_{\tau\text{-}kG,S}$) was the square of the temporal difference between the optimum offset ($k_{G,S}$) and the current offset condition, τ.

The regression equation has five free parameters to be fitted. A single coefficient, α, was used to scale the error term across all conditions and across both groups. Two separate coefficients $\beta_{S=1 \text{ and } 2}$ were used to scale the coordination term for the two sessions; the change in β_S between scanning sessions reflects learning. For the data from the independent condition in experiment 1, we forced these coordination weights β_S to zero. However the weights β_S were fixed across all of the time-offset conditions and across both groups within either session. We also allowed different pre-learning and post-learning constant terms for the two sessions ($\gamma_{S=1,2}$) to reflect all other non-specific changes in BOLD activation between sessions.

Thus, we attempted to fit the mean BOLD signal measured under each experimental condition and in each session by a regression equation based on the measured tracking performance as one regressor (the error term, E, weighted by the single coefficient α) and on the quadratic coordination term C^2 as a second regressor. The latter considers the difference between the current tracking condition and the optimum tracking condition for that subject group, and has a different coefficient from sessions 1 to 2, reflecting learning-related changes in functional activation.

Results

Tracking performance

When first tested in the scanner, both groups of subjects showed the expected quadratic pattern of tracking errors across the five eye–hand temporal offset conditions (experiments 1 and 2, black bars, Figs. 2 and 3A), with better performance in low-asynchrony conditions and higher errors in the higher asynchrony conditions (Miall et al. 2001; Miall and Reckess 2002). In addition, errors were highest in the difficult, independent eye–hand condition, which was tested in group A only (black bars, Fig. 2). There was no significant difference between the tracking errors for the two subject groups in this initial session, tested with a 2×5 (group by condition) repeated measures ANOVA (rmANOVA) ($F_{(1,20)} = 2.95$, $P = 0.1$). Nor was there an interaction between conditions and group ($F_{(1.87,37.38)} = 1.174$, $P = 0.32$, with Greenhouse–Geisser-adjusted d.f).

By fitting quadratic regression curves to the five temporal offset data values, the point of optimum performance was estimated to be when the ocular target led the manual target by $k_1 = 95$ ms (group A) or 83 ms (group B), close to previous values (Miall et al. 2001; Miall and Reckess 2002). The fit of the quadratic curves was high ($r^2 > r^2 > 0.93$, group A: $F_{(1,9)} = 6.96$, $P = 0.027$; group B: $F_{(1,11)} = 16.6$, $P = 0.002$). Thus both the error scores and the point of best performance for the two subject groups were initially comparable.

During training, tracking errors declined progressively over the first 3 days for both groups (Fig. 4). Performance was stable on the fourth, final day of training. Note that the errors in group A are higher than in group B, because the task in which they trained was more difficult, but the learning curves are otherwise similar.

When the two groups were retested after extensive training, there was both an overall reduction in error and a group-dependent shift in the point of best performance (grey bars, Figs. 2 and 3A). A 2×5×2 (group by condition by session) rmANOVA showed there was a

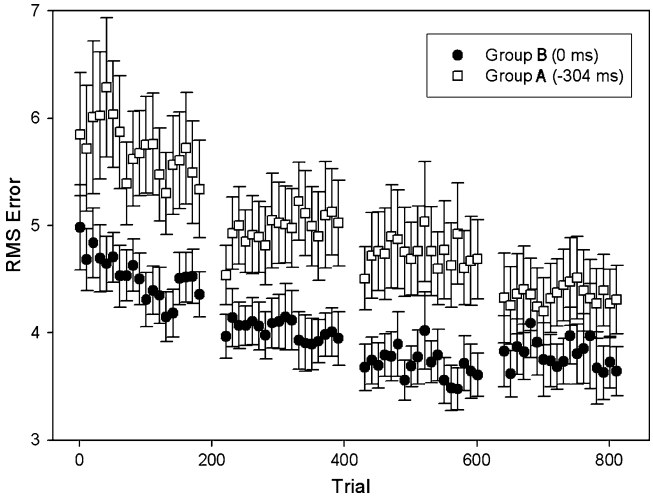

Fig. 4 Learning curves for groups A and B, across 4 days of tracking practice outside the MR scanner. Mean RMS errors have been calculated per subject for blocks of ten trials, and averaged (±1 SEM) across subjects. These curves show the slow decline in error as subjects trained in the task for 1 h per day; performance on the final day seems constant

significant effect of session and condition, as expected but also, critically, there was a significant interaction between group, session, and condition ($F_{(4,80)} = 2.905$, $P = 0.027$). Thus the two groups performed differently across the five conditions, after training in their specific condition.

Importantly, the estimated optimum eye–hand asynchrony point for the two groups shifted in opposite directions upon training. For group A, with subjects trained in the $\tau = -304$ ms condition, it moved from $k_1 = 95$ ms to $k_2 - 77$ ms, i.e. from eye–hand anticipation to hand–eye anticipation. For group B, with subjects trained at the synchronous condition ($\tau = 0$ ms), it moved from $k_1 = 83$ ms to $k_2 = 156$ ms, i.e. toward even greater eye–hand anticipation. Again the fit of the quadratic curves was good ($r^2 > 0.96$ group A: $F_{(1,9)} = 21.25$, $P = 0.001$; group B: $F_{(1,11)} = 5.42$, $P = 0.04$).

It is worth mentioning that part of the overall, nonspecific, reduction in error from sessions 1 to 2, seen in both groups, is probably because of repeated experience of the unusual conditions of the MR magnet. For control groups tested in the laboratory under otherwise similar training conditions smaller mean errors were observed in session 1, and a smaller mean drop in error after training (unpublished work).

Although we did not record eye movements for either subject group, eye movements were recorded outside the scanner using an ASL501 infrared eye tracking system in a separate group of 13 subjects performing the same task used here (five temporal offsets of ±304, ±152, and 0 ms), after training for approximately 30 min. For this group, the median time that the eye was more than 1.5° away from the ocular target was 2.14% of the total. This figure includes all ocular tracking errors, occasional and brief fixations of the cursor, and any apparent error due to drift of the eyetracker calibration; blinks added a further 0.5%. The mean correlation between eye position and target position was $r^2 = 0.89$ (±0.016 SEM); the mean lag of the eye behind the ocular target was 28 ms (±7.8 SEM). Thus we expect that the subjects reported in this paper showed similar accurate and persistent ocular tracking, without significant periods spent fixating the manual cursor or the crosshairs.

For nine subjects from group B who were also tested under experiment 3, the on-line clamping algorithm that adjusted target speeds trial-by-trial to maintain constant RMS error successfully kept mean performance within 5%, both across the two conditions and between sessions 1 and 2 (Fig. 3B). Across all nine subjects, the standard deviation of errors across blocks was only 12% of the target level; within-subject difference in mean error between sessions 1 and 2 was 8% of the target level. Hence a 2×2 (session×condition) rmANOVA showed no significant effect of training, of condition, or an interaction ($F_{(1,9)} < 2.17$, $P > 0.1$). In this experiment, we could, therefore, assess the change in functional activation between the two sessions, free of confounds because of changes in performance.

Functional activation

Training related differences

As an initial analysis we looked for areas activated in common across sessions 1 and 2. There was a highly consistent pattern of activation seen in the comparison of all tracking conditions versus rest for experiments 1–3, with activation of the major sensorimotor areas including strong activation of ipsilateral cerebellum (only data for experiment 2 are presented: Table 1). Next, contrasting the common activation seen across all tracking conditions versus rest between sessions 1 and 2 (pre/post-training) showed no significant *increases* in activity as a result of training in either experiment 1 or 2. There was significantly *reduced* mean activation across all conditions in right sensorimotor areas after training in experiment 1, and in parietal and precuneate areas in

Table 1 Cluster statistics and local maxima for regions showing activation across all five tracking conditions, in sessions 1 and 2

Cluster	Volume (cm^3)	Cluster P	*Z* score	*X* (mm)	*Y* (mm)	*Z* (mm)	Location
1	208.9	0	7.05	44	−62	4	R Middle temporal gyrus, MT/V5
	–		6.34	−44	−70	2	L Middle occipital gyrus, MT/V5
	–		6.61	−32	−22	66	L Precentral gyrus, BA 4
	–		6.41	30	−6	60	R Middle frontal gyrus, BA 6
2	40.1	3.21E−17	5.60	2	−70	−26	R CB, Post. Vermis VII
			5.24	4	−56	−30	R CB, Post. Vermis VII
	–		5.49	28	−46	−36	R CB, Ant. lobule VI simplex
	–		5.24	14	−52	−22	R CB, Ant. lobule V culmen
	–		5.34	16	−56	−30	R CB, dentate nucleus
3	8.23	3.67E−05	5.20	20	−58	54	R Parietal lobe precuneus, BA 7
	–		4.83	16	−62	64	R Superior parietal lobule, BA 7
4	5.4	0.0014	3.85	66	−10	−24	R Inferior temporal gyrus, BA 21
	–		3.29	56	−20	−28	R Fusiform gyrus, BA 20
5	4.1	0.00853	3.38	−42	50	−16	L Middle frontal gyrus
	–		3.33	−34	22	−26	L Superior temporal gyrus
	–		3.10	−40	30	−22	L inferior frontal gyrus

This data are from group B, tested in experiment 2

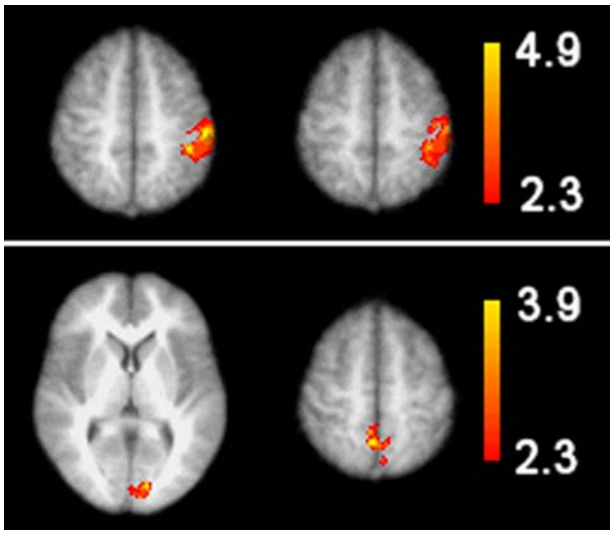

Fig. 5 Areas of significant decrease in mean BOLD signal after training: *top*, experiment 1; *bottom*, experiment 2. Each panel shows selected horizontal slices through the regions of reduced activity, contrasting all active tracking conditions from session 1 against session 2. Table 2 provides details of loci; colour bars indicated *Z* scores. Horizontal slices in the *top panel* are at +52 and +56 mm above the origin; slices in the *bottom panel* are at +4 and +58 mm

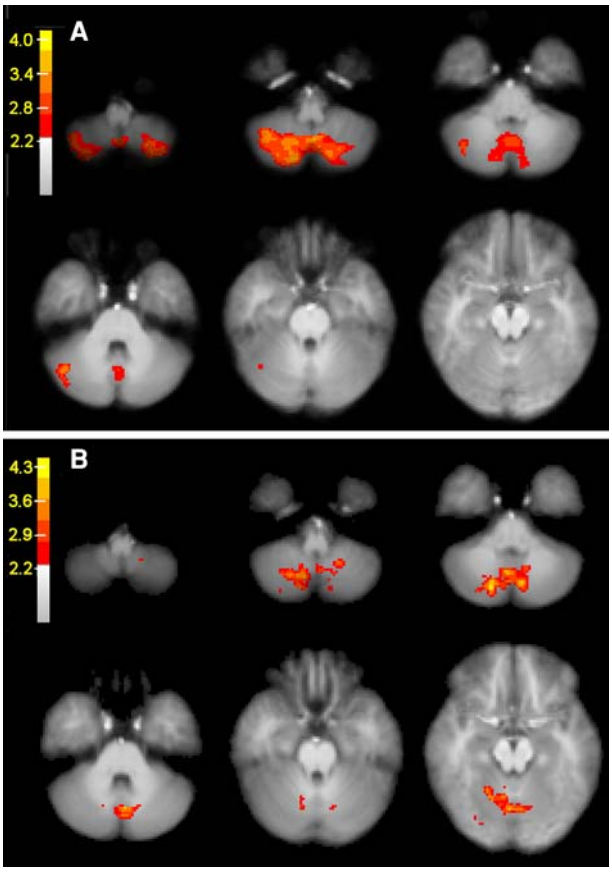

Fig. 6 Areas of significant increase in quadratic regression weights between tracking condition and the BOLD signal after training. **A** experiment 1. **B** experiment 2. Each panel shows the region of activity within the cerebellum, contrasting the non-linear regression model from session 2 against session 1 (Eq. 3). Table 3 provides details of loci; *colour bars* indicate *Z* scores. Horizontal slices in **A** are at 8 mm separation from −60 to −20 mm; slices in **B** are at 8 mm separation from −58 to −18 mm

experiment 2 (Fig. 5, Table 2). Pair-wise comparison of each tracking condition before and after training showed reduced activation of sensorimotor areas in the two most extreme temporal offset conditions ($\tau = \pm 304$ ms), but the other within-condition and between-session differences did not reach significance.

In summary, the expected sensory-motor areas of the cerebrum and cerebellum were strongly activated during the tracking task, and there was a generalised decrease in activity in these areas after training, with a greater decrease during the more difficult conditions that had large eye–hand temporal offsets.

Time-offset-related activation differences

We next tested across the whole brain for areas with differential parametric activation after training. In other words, we contrasted the slope of the parametric fits between BOLD signal and tracking condition, before and after training. This analysis identifies those areas in which the slope of the quadratic curve relating time-offset to BOLD signal was higher in session 2 after training than in session 1, i.e. where:

Table 2 Cluster statistics and local maxima for areas showing reduced activation across all tracking conditions, from sessions 1 and 2

Cluster	Vol. (cm^3)	Cluster Level P	*Z* score	*X* (mm)	*Y* (mm)	*Z* (mm)	Location
EXP 1							**Session 1 > session 2**
1	11.8	0.00282	4.91	−54	−24	52	L Postcentral gyrus, BA 2
	–	–	4.65	−42	−38	52	L Inferior parietal lobule, BA 40
	–	–	3.54	−46	−20	56	L Postcentral gyrus, BA 3
EXP 2							**Session 1 > session 2**
1	3.1	0.00348	3.74	−12	−84	4	L Cuneus, BA 17
	–	–	3.29	10	−86	10	R Cuneus, BA 17
2	2.9	0.00497	3.94	6	−56	58	R Precuneus, BA7
	–	–	3.12	10	−54	70	R Postcentral gyrus, BA 7
	–	–	3.05	−4	−70	58	L Superior parietal lobule, BA 7

Table 3 Cluster statistics and local maxima of areas with increased weighting of the quadratic regression of BOLD signal and temporal offset condition, for session 2 compared with session 1

Cluster	Vol. (cm^3)	Cluster Level P	*Z* score	*X* (mm)	*Y* (mm)	*Z* (mm)	Location
EXP 1							**Session2 > session1**
1	26.9	0.000265	3.90	−22	68	22	L Superior frontal gyrus, BA 10
	–	–	3.41	−10	70	4	L Superior frontal gyrus, BA 10
	–	–	3.53	−28	62	26	L Middle frontal gyrus, BA 8
	–	–	3.47	16	30	60	R Superior frontal gyrus, BA 10
2	24.4	0.000548	3.43	−2	−58	−54	L CB, Post. Vermis Lobule VIII
	–	–	3.36	40	−52	−54	R CB, Post. lobule VII paramedian
	–	–	3.29	16	−74	−52	R CB, Post. lobule VIII biventer
EXP 2							**Session 2 > session 1**
1	15.2	2.38E−07	4.26	-6	−68	−40	L CB, Post. lobule VIII biventer
	–		4.00	18	−74	−42	R CB, Post. lobule crus II
	–		3.87	10	−60	−46	R CB, Post. lobule VIII paramedian
	–		3.84	4	−60	−42	R CB, Post. vermis VIII
	–		3.80	14	−70	−44	R CB, Post. lobule VIIB/Crus II
EXP 3							**Session 2 > session 1**
1	2.9	0.00134	4.11	−4	−76	−40	L CB Posterior vermis VIIA
	–	–	3.98	−12	−82	−42	L CB Posterior lobule crus II
	–	–	3.77	−26	−86	−38	L CB Posterior lobule crus I/II
	–	–	3.17	2	−84	−44	R CB Posterior lobule crus II

$$CW \times PE_{\text{session 2}} - CW \times PE_{\text{session 1}} > 0.0 \quad (3)$$

We found areas within cerebellum had significantly increased slope in both experiments (Fig. 6), with local maxima found in both posterior cerebellar lobules and vermis (Table 3). Note that by testing for this increased slope across the tracking conditions we could identify areas as having a higher slope in session 2 (Table 3), even though the *average* level of BOLD signal across all tracking conditions was lower in session 2 than in session 1 (cf. Fig. 5). In experiment 1, a region of left frontal pole/prefrontal cortex was also identified; in experiment 2 a small number of voxels in the right occipital cortex were also significantly different, close to but not overlapping an area of strong activation seen in the main tracking conditions (area MT/V5; Table 1; see bottom right slice in Fig. 6B). No brain areas showed significant *reduction* in slope of the regression model (i.e. Eq. (1) < 1.0) after training for either experiment.

Region-of-interest analysis

To further quantify these activation patterns, we measured the mean percentage BOLD signal levels for each task condition in experiment 1 (i.e. the mean values of the six PEs from the general linear modelling of BOLD signal), using as region of interest the volume of significant cerebellar activation shown in Fig. 6A. Within this volume there was high activation in the first session, in all the conditions with high tracking errors (independent and

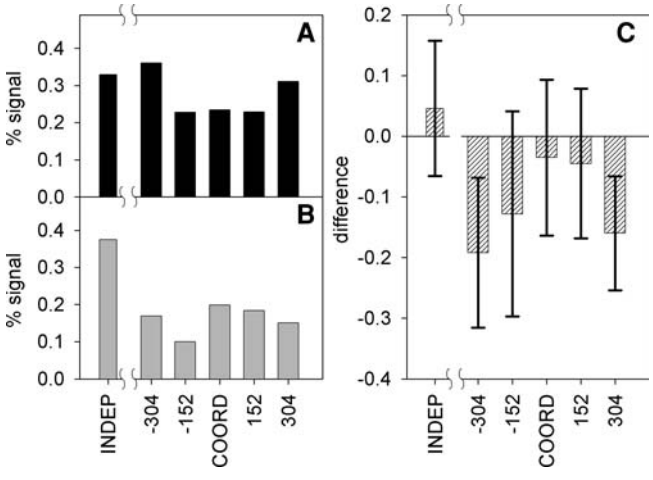

Fig. 7 ROI analysis of BOLD signal levels, experiment 1. **A**. Mean GLM PEs (*bars*, see "Methods") were measured during session 1 from the ROI shown in Fig. 6A (volume: 185 ± 6 voxels) **B**. Mean PEs in session 2. **C**. Within-subject difference in mean PEs (± SEM, $n = 10$), sessions 2 to 1

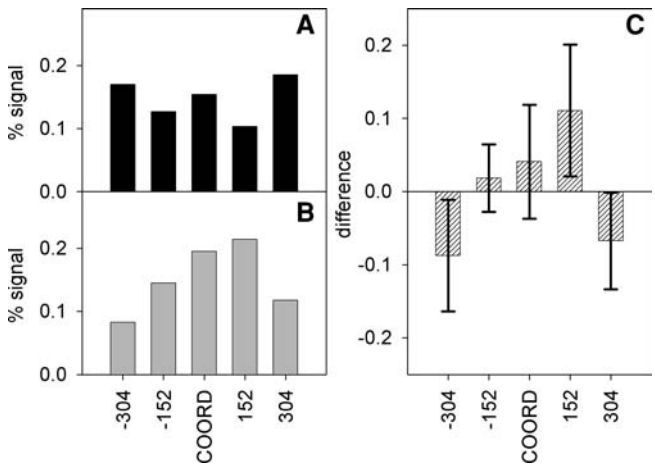

Fig. 8 ROI analysis of BOLD signal levels, experiment 2. **A**. Mean PEs (*bars*, see "Methods") were measured during session 1 from the ROI shown in Fig. 6B (volume: 97 ± 4 voxels). **B**. Mean PEs in session 2. **C**. Within-subject difference in mean PEs (± SEM, $n = 12$), sessions 2 to 1

$\tau = \pm 304$ ms temporal offset; Fig. 7A). In the second session, there was overall decline in mean activation (Fig 7A compared with Fig 7B), combined with development of the expected non-linear relationship between the tracking condition and BOLD signal. Hence, a shallow quadratic relationship between BOLD and eye–hand asynchrony present in session 1 (Fig. 7A) was replaced by a steeper and inverse quadratic relationship in session 2 (Fig. 7B). The between-sessions change in signal within this region, measured within-subject, confirmed this, showing a strong inverse quadratic relationship across the five temporal offset conditions (Fig. 7C).

A similar pattern of activation was seen across the five conditions of experiment 2, using the cerebellar areas identified in Fig. 6B as a region of interest. In the first session (Fig. 8A) there was again an increased signal at the $\tau = \pm 304$ ms temporal offset conditions, when tracking errors were elevated (Fig. 3A). After training, a higher signal was then seen in the coordinated ($\tau = 0$ ms) and $\tau = +152$ ms conditions (Fig. 8B), when performance was best and errors least. The within-subject difference in BOLD activation from sessions 1 to 2 confirmed this pattern, with clear inverse U-shaped curve being seen (Fig. 8C).

Fitting quadratic curves to the BOLD signal across the five temporal asynchrony conditions from the pre-training session for experiments 1 and 2, we estimated the inflexion point in the BOLD curve to be at $+25$ and -4 ms, respectively (cf. Figs. 7A and 8A). After training, the inflexion points moved to -122 and $+42$ ms, respectively.

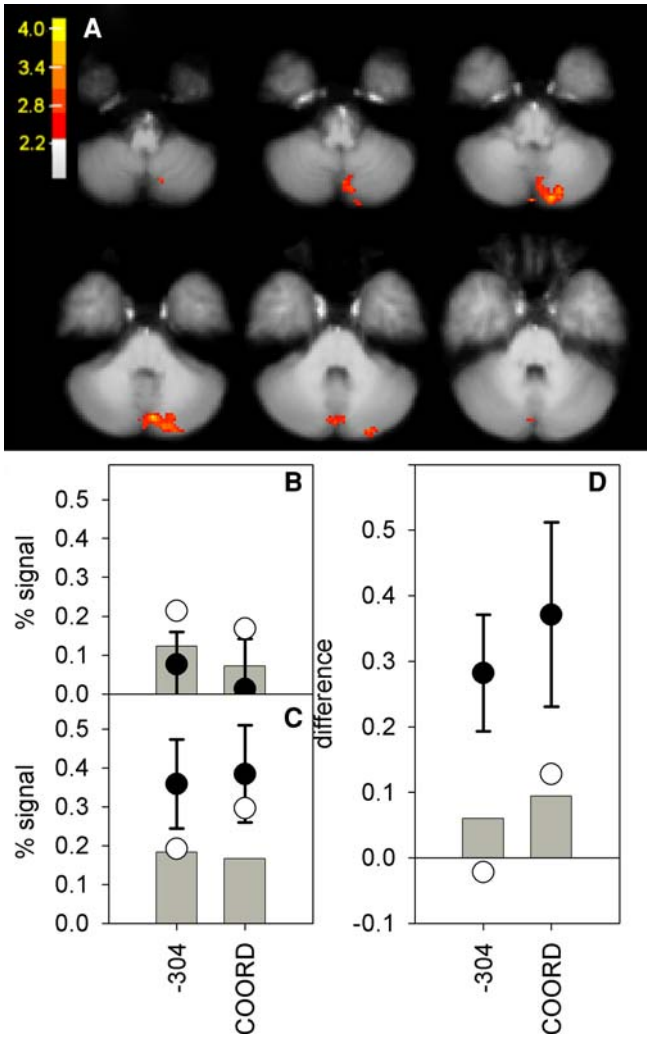

Fig. 9 Performance-independent activation changes, experiment 3. **A**. The areas within the cerebellum showing a significant increase in activation during session 2, despite constant performance across sessions (Fig. 3B). **B**. Mean PEs (*grey bars*) measured during session 1 from an ROI analysis using the cluster shown in Fig. 9A (volume: 17 ± 1 voxels). *Black dots* represent mean PEs ($\pm$SEM, $n = 9$) calculated using the coordinate of maximum significance shown in panel A (*XYZ*: $+44{:}-76{:}-40$ mm). *White dots* show the predicted PEs (see "Discussion"). **C**. Mean PEs, session 2; same format as panel **B**. **D**. Within-subject differences in mean PEs, sessions 2 to 1; same format as panel **B** (*grey bars*). *White dots* are the mean differences predicted from Eq. (2). *Black dots* are the average within-subject differences (± 1 SEM) at the coordinate of maximum significance

Performance-independent activation differences

In experiment 3 we controlled target speeds, to clamp tracking performance at a constant level between sessions 1 and 2 (Fig. 3B). Despite this constant level of performance, there was a significant increase in activation within the cerebellum after training (Fig. 9A). Measuring mean BOLD signal levels within this volume of significance showed an increase in activation in both tracking conditions in session 2 (Fig. 9C compared with Fig. 9B). The within-subject mean activation differences

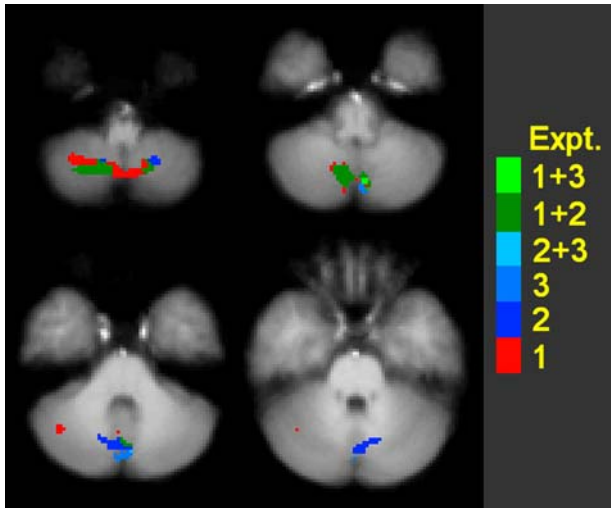

Fig. 10 Spatial overlap between areas activated during tracking and areas with significant change in activation after learning. Activation clusters for all three experiments (Figs. 6A, B and 9A) were masked by the area of activation found across all five tracking conditions, sessions 1 and 2, experiment 2 (Table 1). Regions of overlap are colour coded by experiment number. Horizontal slices are at -54, -46, -38 and -30 mm

measured between sessions 1 and 2 tended towards a greater increase for the trained, coordinated, condition than for the $\tau = -304$ ms temporal offset condition (Fig. 9D). The region of significant activation shown in Fig. 9A is close to but not identical to that seen in Fig. 6B. We therefore repeated the ROI analysis shown in Fig. 9D using a region specified by the activation cluster from Fig. 6B; the results were highly comparable: differences from the data shown in Fig. 9D were under 5% (difference in % BOLD signal of $< 0.004\%$).

Comparison of spatial activation patterns

Figures 6A, B suggest that the areas in the cerebellum that increased activation patterns after training depend on the training condition used. Figure 10 shows the overlap between activation maps for all three experiments (Figs. 6A, B and 9A), masked by the region activated in common across all tracking conditions (Table 1). This emphasises that the areas showing a change in activity upon learning have considerable spatial overlap with the areas activated by the tracking task itself. Moreover, two regions showed the strongest overlap: the paramedian and biventer lobules (VIII) and the vermis (VII). For group A (experiment 1), the overlap was mainly in the biventer lobule (Fig. 10, red and green areas). Remember that this group were trained in the condition where the manual target preceded the ocular target, and they shifted their point of peak performance—i.e. the minimum of the error curves shown in Fig. 2—from $k_1 = +95$ ms (eye leading hand) to $k_2 = -77$ ms (hand leading eye). For group B (experiment 2), who where trained in the null, coordinated condition, the areas of functional activation overlap tested in experiments 2 and 3 were largely in the ocular vermis (Fig. 10, blue areas). This group shifted their peak performance toward even higher ocular lead times, from $k_1 = 83$ ms to $k_2 = 156$ ms. Finally, there was also some common activation across the two subject groups, in the biventer lobule (Fig. 10, green).

Discussion

Manual tracking performance during a dual ocular and manual tracking task was previously reported to depend on the degree of temporal coupling between the eye and hand movements (Miall and Reckess 2002). We show here that this behaviour is sensitive to training, and that performance changes differentially with experience of different eye–hand temporal offsets. We also show that patterns of cerebellar functional activation which have been previously reported (Miall et al. 2001) also change with training, and this change is consistent with the cerebellum providing experience-dependent coordination of eye and hand actions.

Estimating cerebellar BOLD signal

We propose that the cerebellar functional activation observed with fMRI in this task is related to two performance factors. First, there is a contribution directly related to the level of tracking error (Imamizu et al. 2000; Miall et al. 2001; Nezafat et al. 2001), and this caused high overall BOLD signal in the first session, with the highest activation under the most difficult conditions—the extreme temporal offset conditions (± 304 ms) and the independent condition of experiment 1. This pattern is clearly seen in Fig. 7A. After training, performance improved across all tracking conditions and, therefore, this component contributed less to the BOLD signal in session 2 than in session 1. Thus the

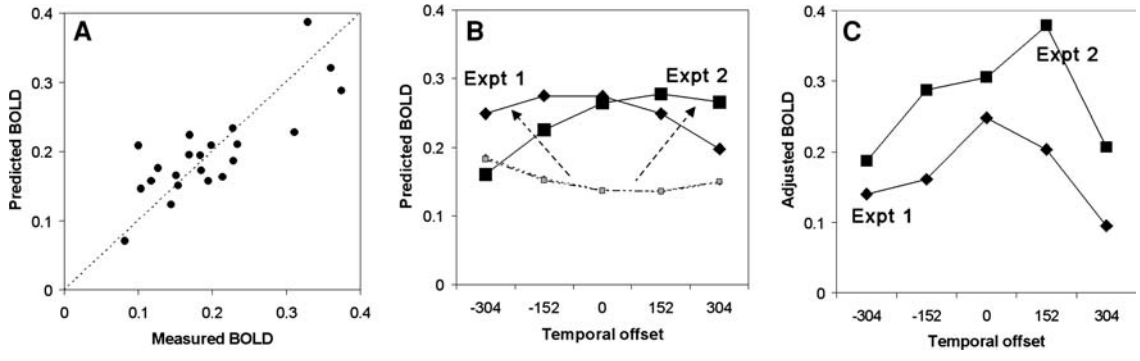

Fig. 11 Regression analysis of the BOLD signal. **A**. The mean BOLD signal measured in each condition and session for experiments 1 and 2 ($n = 22$) plotted against the signal predicted by the regression model, Eq. (2). **B**. The BOLD signal predicted from Eq. (2), after subtraction of the estimated contributions of the error (α) and constant components (γ). The predicted signal during session 1, across the five temporal offset conditions, was almost identical for experiments 1 and 2 (*dotted lines*). In session 2, the quadratic relationship was inverted, and shifted toward positive (experiment 2, *squares*) or negative (experiment 1, *diamonds*) time-offsets. The *arrows* indicate the overall shift for each experiment from the initial curve. **C**. The adjusted BOLD signals calculated as the measured differences in BOLD signal between sessions 2 and 1, adjusted by the subtraction of the estimated contributions from change in performance and from the constant components of Eq. (2). These curves show changes in BOLD signal we would expect to measure, had we been able to compensate for performance differences between sessions, and for non-specific reduction in activation signal in session 2 (cf. Figs. 7C and 8C)

average activation across the tracking conditions was lower, although it was still quite high in the independent condition in which errors were still high (Fig. 7B). This pattern replicates the results reported previously (Miall et al. 2001).

However, the second contributing factor is, we propose, related to an internal predictive model that develops as a result of the subjects' training and experience at the task and this causes a non-linear, inverse quadratic relationship between temporal offset condition and BOLD (Miall et al. 2001). This factor should then contribute most strongly to the measured BOLD signal at the tracking condition in which the subjects were trained, because they will gain most experience in that condition. It should contribute less at conditions further from the trained condition. This hypothesis predicts an inverse quadratic relationship between tracking condition and BOLD signal, with the maximum BOLD signal shifted toward the training condition for each subject group. It also predicts that there should be an increase in BOLD signal after training even under error-clamped conditions, as we have found in experiment 3. Last, the estimated point of best tracking performance, which moved in opposite directions from sessions 1 to 2 in the two groups, should be close to the estimated maximum of the curve fitted to the BOLD signal. This was reasonably well matched: the estimated inflexion point of the quadratic BOLD curves shifted by −147 ms in group A and by +46 ms in group B, while their peak performance shifted by −172 ms and +74 ms, respectively.

We therefore tested the predictive power of this relationship between performance and BOLD signal recorded under all conditions in the four experimental sessions ($n = 22$, experiments 1 and 2, sessions 1 and 2). The regression equation (''Methods'', Eq. 2) used the group mean tracking error scores for each condition ($E_{G,S}$, Figs.2 and 3) as a linear regressor. The quadratic component (the C^2 term in Eq. 2) used the squared differences between each temporal offset condition (τ) and the point of optimum tracking (k_S), the latter also estimated from the performance curves measured during each experiment. The weighting coefficient for this quadratic term, β_S, was allowed to vary between pre-learning and post-learning sessions to reflect the learning process. Finally we used two constants, γ_S, for sessions 1 and 2, to account for non-specific changes in BOLD activation between sessions. Thus all regression predictors were derived from measured performance curves and there were five free parameters. The regression was highly significant ($r^2 = 0.66$, $P < 0.001$, $n = 22$, Fig. 11A). The coefficient for the quadratic component rose from $\beta_1 = -0.32$ to $\beta_2 = 0.55$ between sessions 1 and 2, confirming an increased contribution to the BOLD signal from the second, coordination-related factor, as a result of learning. We then used these same regression coefficients measured, under experiments 1 and 2, to predict the BOLD signal expected in experiment 3, which had been performed under constant error conditions. The predictions were close to the mean BOLD values measured, and the relative changes between the two conditions and across the two sessions are in the appropriate directions (Fig. 9, white dots).

Our regression model assumes that the error component is constant before and after learning, as only a single coefficient, α, was used. It has been suggested that this assumption may be wrong, and that the error term might also change as function of training. Hence, an alternative regression model was tested with separate error terms for sessions 1 and 2, thus increasing the number of free parameters, but this resulted in an equally good fit to the data ($r^2 = 0.66$). The two error coefficients in the new model were within 8 and 25% of the original single error term. More importantly, the coefficients for the coordination term, β were almost identical across these two regression models—they changed by less than 2%. In both models, all regression coefficients except $\beta_1$were significantly different from zero. Thus either regression model is robust, and they only differ in their parcellation of weights between the error term α and the constant terms γ_S from sessions 1 to 2. Both models, however, suggest the coordination term becomes significant after learning.

Decomposing the BOLD signal

The good fit of the regression model (Eq. (2), Fig. 11A) supports our suggestion that performance error and a coordination factor both add to the measured BOLD signal. Hence, the BOLD levels seen in sessions 1 and 2 reflect contributions from both terms. The regression equation can be used to estimate the BOLD signals that might be expected in ideal conditions, uncontaminated by any performance changes, by setting the error coefficients to zero and adjusting for the constant terms (Fig. 11B). Alternatively, using the error regression coefficient ($\alpha = 0.0742$) we can estimate and remove from the measured BOLD signal any contribution from these terms, leaving an adjusted BOLD estimate (Fig. 11C). Note that the curves in Fig. 11C are similar to those in Figs. 7C and 8C, shifted positively to account for the drop in performance errors across sessions. From either method we would expect an increase in BOLD signal in the cerebellum after learning, if the performance errors did not change. This was confirmed directly by experiment 3.

Specificity of learning

Optimum manual tracking performance in this dual eye and hand tracking task is normally found when the ocular target leads the manual target by $k = 75$–90 ms (Figs. 2 and 3, black bars, and Miall et al. 2001; Miall and Reckess 2002). Ocular performance shows only modest changes across conditions in this task, and the eyes faithfully track the ocular targets for the great majority of every trial (unpublished data). Hence suc-

cessful oculomotor tracking of the ocular target is probably used to derive a predictive signal that then assists manual tracking, and this process is improved by practice at the task. In experiment 2, the additional learning experience in the synchronous condition ($\tau = 0$ ms) shifted this group's point of optimum performance to $k_2 = 156$ ms, i.e. to a point at which the best performance was seen when the eye movement preceded the hand by 156 ms. This shift may reflect a balance between the extent to which longer lead times gained from eye movement can assist in predictive programming of arm movements, and an upper useful limit of the associations between the two signals. Clearly if the eye anticipated the hand by many hundreds of milliseconds, the predictive value would be low.

However, manual-tracking performance was also improved for group A, who trained under the condition when the ocular target lagged behind the manual target, and in this case the predictive information from ocular pursuit of the target should be of limited utility. So in this case, the subjects may have partly decoupled the now inappropriate ocular signals from the manual control signals, to improve manual performance. This suggests that experience of different eye–hand temporal coupling conditions alters the balance between the relative contribution of manual and ocular control signals. Although task-specific learning and thus a shift of peak performance toward their training condition of −304 ms was observed for subjects in experiment 1, we do not yet know if subjects would shift even further toward the −304 ms point if they were given more extended training, nor whether they could learn under more extreme temporal offsets. By analogy with the predictive interval between conditioned stimuli and responses in associative learning tasks (Ohyama et al. 2003), we would expect greater difficulty at longer intervals.

The change in BOLD signal seen within the cerebellum further suggests that it is the neural locus of the practice-related learning. Imamizu et al. (2000, 2003) argue for tool-specific localised changes in cerebellar cortical activation; activation in their studies is more lateral than ours, in areas reported active in cognitive tasks or attention demanding tasks (Allen et al. 1997; Cabeza and Nyberg 2000). In our task, learning does not involve adapting to new or modified tools, but adaptation to changed temporal relationships between eye and hand. That our subjects are not learning a novel "tool" is further supported by finding that the areas of significant change in activation after learning are indeed visuomotor areas (Fig. 5). In fact, we found spatial overlap in all three experiments between those areas activated due to the basic tracking conditions and the areas showing significant change in activation after learning (Fig. 10). Two regions showed greatest overlap, the paramedian and biventer lobules (VIII; Schmahmann et al. 1999) and the vermis (VII). Moreover, the overlap seems to depend on the training conditions. In experiment 1, subjects were trained in the condition where the manual target anticipated the ocular target, and shifted their peak performance accordingly (hand leading eye). In this experiment, the spatial overlap between those voxels showing significant change in functional activity and voxel active across all tracking conditions was mainly in the biventer lobule. For experiment 2, subjects trained in the coordinated condition and shifted their peak performance toward even higher ocular lead times. The areas of activation overlap for this group (experiments 2 and 3) were largely in the ocular vermis.

This pattern suggests that local visuo-motor areas within cerebellar cortex are differentially affected by learning. When the hand action predicts ocular motion, an increased component of the BOLD signal related to the time-offsets is observed in paramedian/biventer areas related to limb movement (Grafton et al. 1992; Miall et al. 2000, 2001). In contrast, when ocular motion predicts hand action, an increased component of the BOLD signal related to the time-offsets is observed in the oculomotor vermis (Lewis and Zee 1993; Petit et al. 1996; Carter and Zee 1997; Desmurget et al. 1998; Miall et al. 2000).

Finally, we note that (Seidler et al. 2002) recently argued that the cerebellum was not the locus for motor learning, because they saw no activation differences when performance was held constant, despite learning being demonstrated. Their task was, however, the serial reaction time task in which an important aspect of learning is the implicit learning of a sequence of button press responses. The difference between our tasks implies that sequence learning may involve areas other than the cerebellum, but for learned eye–hand coordination the cerebellum is a major locus supporting learning.

In summary, we suggest that the cerebellar activation we have observed across the different eye–hand time-offset conditions has contributions related to performance error and to the expected temporal relationships between eye and hand movement. The latter component changes after a period of training, consistent with the observed change in performance. We suggest that this cerebellar activity represents a predictive signal, originating from cerebellar areas concerned with control of the leading effector (eye or hand), which is used to improve the oculo-manual tracking performance. In sum, the cerebellar processing represents the internal forward model signal that we propose underlies coordinated action, combined with an error term that might itself be used to drive learning of the forward model.

Acknowledgements This work was funded by a Wellcome Trust Fellowship to RCM and by MRC support of the Oxford FMRIB Centre. We thank all FMRIB staff for their technical support, analysis software and advice, and thank Ji-Hang Lee for his data on eye tracking performance mentioned in the results.

References

Allen G, Buxton RB, Wong EC, Courchesne E (1997) Attentional activation of the cerebellum independent of motor involvement. Science 275:1940–1943

Cabeza R, Nyberg L (2000) Imaging cognition II: an empirical review of 275 PET and fMRI studies. J Cogn Neurosci 12:1–47
Carter N, Zee DS (1997) The anatomical localization of saccades using functional imaging studies and transcranial magnetic stimulation. Curr Opin Neurol 10:10–17
Desmurget M, Pelisson D, Urquizar C, Prablanc C, Alexander GE, Grafton ST (1998) Functional anatomy of saccadic adaptation in humans. Nat Neurosci 1:524–528
Duvernoy HM (1999) The human brain surface, blood supply, and three-dimensional sectional anatomy. Springer-Verlag Wein, New York
Forman SD, Cohen JD, Fitzgerald M, Eddy WF, Mintun MA, Noll DC (1995) Improved assessment of significant activation in functional magnetic resonance imaging (fMRI): use of a cluster-size threshold. Magn Reson Med 33:636–647
Friston KJ, Worsley KJ, Frackowiak RSJ, Mazziota JC, Evans AC (1994) Assessing the significance of focal activations using their spatial extent. Hum Brain Mapp 1:214–220
Grafton ST, Mazziotta JC, Presty S, Friston KJ, Frackowiak RSJ, Phelps ME (1992) Functional anatomy of human procedural learning determined with regional cerebral blood flow and PET. J Neurosci 12:2542–2548
Imamizu H, Miyauchi S, Tamada T, Sasaki Y, Takino R, Putz B, Yoshioka T, Kawato M (2000) Human cerebellar activity reflecting an acquired internal model of a new tool. Nature 403:192–195
Imamizu H, Kuroda T, Miyauchi S, Yoshioka T, Kawato M (2003) Modular organization of internal models of tools in the human cerebellum. Proc Natl Acad Sci USA 100:5461–5466
Ito M (2002) Historical review of the significance of the cerebellum and the role of Purkinje cells in motor learning. Ann NY Acad Sci 978:273–288
Jenkins IH, Brooks DJ, Nixon PD, Frackowiak RSJ, Passingham RE (1994) Motor sequence learning: a study with positron emission tomography. J Neurosci 14:3775–3790
Jenkinson M, Smith S (2001) A global optimisation method for robust affine registration of brain images. Med Image Anal 5:143–156
Jenkinson M, Bannister P, Brady M, Smith S (2002) Improved optimization for the robust and accurate linear registration and motion correction of brain images. Neuroimage 17:825–841
Jueptner M, Frith CD, Brooks DJ, Frackowiak RSJ, Passingham RE (1997a) Anatomy of motor learning 2 Subcortical structures and learning by trial and error. J Neurophysiol 77:1325–1337
Jueptner M, Stephan KM, Frith CD, Brooks DJ, Frackowiak RSJ, Passingham RE (1997b) Anatomy of motor learning 1 Frontal cortex and attention to action. J Neurophysiol 77:1313–1324
Lewis RF, Zee DS (1993) Ocular motor disorders associated with cerebellar lesions: pathophysiology and topical localization. Rev Neurol (Paris) 149:665–677
Miall RC, Reckess GZ (2002) The cerebellum and the timing of coordinated eye and hand tracking. Brain Cogn 48:212–226
Miall RC, Imamizu H, Miyauchi S (2000) Activation of the cerebellum in co-ordinated eye and hand tracking movements: an fMRI study. Exp Brain Res 135:22–33
Miall RC, Reckess GZ, Imamizu H (2001) The cerebellum coordinates eye and hand tracking movements. Natl Neurosci 4:638–644
Nezafat R, Shadmehr R, Holcomb HH (2001) Long-term adaptation to dynamics of reaching movements: a PET study. Exp Brain Res 140:66–76
Ohyama T, Nores WL, Murphy M, Mauk MD (2003) What the cerebellum computes. Trends Neurosci 26:222–227
Petit L, Orssaud C, Tzourio N, Crivello F, Berthoz A, Mazoyer B (1996) Functional anatomy of a prelearned sequence of horizontal saccades in humans. J Neurosci 16:3714–3726
Ramnani N, Passingham RE (2001) Changes in the human brain during rhythm learning. J Cogn Neurosci 13:952–966
Ramnani N, Toni I, Josephs O, Ashburner J, Passingham RE (2000) Learning and expectation-related changes in the human brain during motor learning. J Neurophysiol 84:3026–3035
Schmahmann JD, Doyon J, McDonald D, Holmes C, Lavoie K, Hurwitz AS, Kabani N, Toga A, Evans A, Petrides M (1999) Three-dimensional MRI atlas of the human cerebellum in proportional stereotaxic space. Neuroimage 10:233–260
Seidler RD, Purushotham A, Kim SG, Ugurbil K, Willingham D, Ashe J (2002) Cerebellum activation associated with performance change but not motor learning. Science 296:2043–2046
Shadmehr R, Holcomb HH (1997) Neural correlates of motor memory consolidation. Science 277:821–825
Smith SM (2001) Overview of fMRI analysis. In: Jezzard P, Matthews PM, Smith SM (eds) Functional MRI: an introduction to methods. Oxford University Press, Oxford, pp 215–227
Talairach J, Tournoux P (1988) Co-planar stereotaxic atlas of the human brain. Georg Thieme Verlag, Stuttgart
Tamada T, Miyauchi S, Imamizu H, Yoshioka T, Kawato M (1999) Cerebro-cerebellar functional connectivity revealed by the laterality index in tool-use learning. Neuroreport 10:325–331
Thach WT, Goodkin HP, Keating JG (1992) The cerebellum and the adaptive coordination of movement. Ann Rev Neurosci 15:403–442
Turner RS, Grafton ST, Votaw JR, DeLong MR, Hoffman JM (1998) Motor subcircuits mediating the control of movement velocity: a PET study. J Neurophysiol 80:2162–2176
Woolrich MW, Ripley BD, Brady M, Smith SM (2001) Temporal autocorrelation in univariate linear modeling of FMRI data. Neuroimage 14:1370–1386
Worsley KJ, Evans AC, Marrett S, Neelin P (1992) A three-dimensional statistical analysis for rCBF activation studies in human brain. J Cereb Blood Flow Metab 12:900–918