Supplemental Results

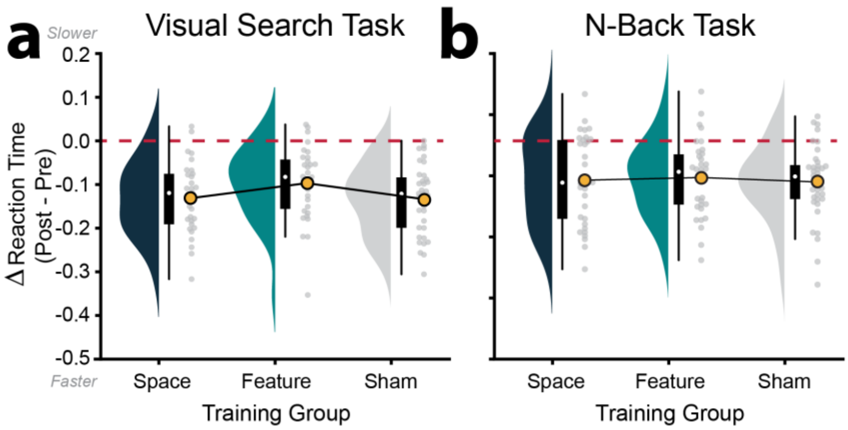
Transfer Tasks

In addition to the visual selective attention test used to assess training effects, which can be considered a near transfer task, participants also performed two additional transfer tasks before and after training. These included a visual search task, on which one might expect some transfer of visual selective attention training effects, and a visual-verbal n-back task, on which training effects specific to visual selective attention would not be expected to transfer.

Visual Search Task. To assess performance on the visual search task, we recorded reaction times to the task for each test day (Before Neurofeedback, After Neurofeedback). Using a pairwise JZS t-test, we found decisive evidence that reaction times on the visual search task improved overall (All Groups BF10 = 1.48x1011; space neurofeedback group: M = -131.91, SD = 102.88; feature neurofeedback group: M = -97.16, SD = 99.42; sham neurofeedback group: M = -133.79, SD = 98.60). The visual search task featured three set-sizes. We therefore submitted reaction time training effects (After – Before Neurofeedback, Δms) to a 2x3 Bayesian ANOVA with Neurofeedback Group (Active, Sham) and set-size (Small, Medium, Large) as factors. We found substantial evidence for a null interaction between these factors ($\text{Interaction}_{\text{NF-typeXset-size}}$: $BF_{10} = 0.16$), and anecdotal evidence that reaction time training effects did not differ across neurofeedback groups (Main Effect $BF_{10} = 0.47$, **Supplemental Figure 1*a***). Thus, there is no evidence that the neurofeedback-specific training effects transferred to this different selective attention task.

N-Back Task. To assess performance on the N-back task, we also recorded reaction times to the task for each test day (Before neurofeedback, After neurofeedback). Using a pairwise JZS-test, we found decisive evidence that reaction times on the N-back task improved overall (All Groups BF10 = 5.56x107; space neurofeedback group: M = -90.35, SD = 101.69; feature neurofeedback group: M = -84.30, SD = 86.10; sham neurofeedback group: M = -94.57, SD = 85.81). Reaction time training effects (After – Before Neurofeedback, Δms) on the N-back task were submitted to a JZS t-test across Neurofeedback Group (Active, Sham). This revealed substantial evidence for a null effect of Neurofeedback Group ($BF_{10} = 0.23$, **Supplemental Figure 1*b***).

Thus, while Neurofeedback conferred a training advantage on a task similar to that used during training (near-transfer) there was no evidence that this advantage transferred to either the visual search attention task or n-back working memory task.



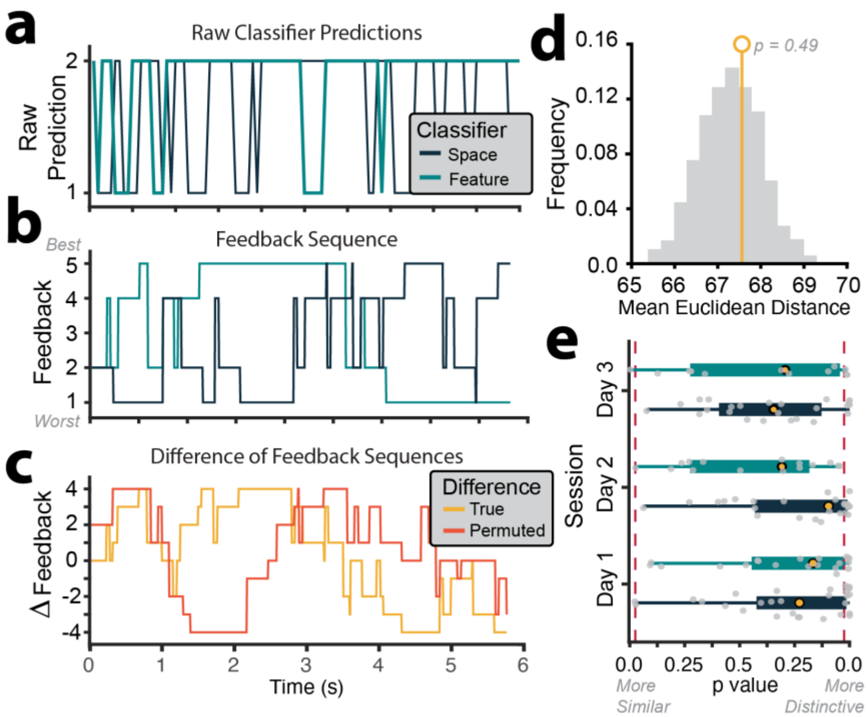
Supplementary Figure 1. Behavioural training effects on the transfer tasks. Raincloud plots show the change in reaction time from the pre- to post- neurofeedback training test of the **a.**) visual search and **b.**) N-back task for each training group.

Feedback Specificity

Behavioural training effects were not specific to the trained mode of attention, suggesting that feedback was broadly integrated. One compelling alternative explanation could be a lack of specificity in the feedback signals. It is possible that spatial and feature-based attention might have been strongly correlated in time during this task. Thus, non-specific training effects may have arisen due to highly correlated feedback signals across the two types of neurofeedback. We assessed this possibility for our training task by comparing the presented feedback with that which would have been presented if the participant were in the other active training condition. Using these data, we asked whether the spatial and feature-based neurofeedback streams were more similar than would be expected by chance.

We performed a permutation test for each neurofeedback session, using the Euclidean distance between the frame-by-frame spatial and feature-based neurofeedback sequences as a measure of similarity (**Supplementary Figure 2a, 2b**). Next, we computed a null distribution (N_{perm} = 1000) of the distances which might be expected for such feedback signals by chance. This was done by re-computing the Euclidean distance using shuffled trial indices such that no trial was matched with itself. For each shuffled trial, we flipped the signal either vertically (i.e. low values become high) or horizontally (i.e. in time) with equal probability. This flipping procedure removes any common temporal patterns in feedback across trials, setting a high bar for feedback differences, thus making it more likely that any real feedback time-courses would be identified as similar to each other (**Supplementary Figure 2c, 2d**).

Despite this high bar, only 2 out of 213 neurofeedback sessions (0.94%) met the threshold for significance (α = .05, two tailed: *sub-28 feature-based training day 1, p = .01 | sub-72 spatial training day 3, p = .001,* **Supplementary Figure 2e**). These sessions can likely be attributed to type I error. No other sessions and no other subjects met the threshold for significant similarity between feature-based and spatial attention neurofeedback signals. Thus, each type of active neurofeedback generated unique feedback signals, and we can rule out the possibility that any non-specific training effects might have arisen solely from non-specific feedback.



Supplementary Figure 2. Results of the permutation test conducted to assess the independence of neurofeedback signals generated for each active neurofeedback type (Space, Feature). **a.)** We trained both a space and a feature-neurofeedback classifier for each active neurofeedback participant, though only one of these was used to generate feedback in the actual experiment. For this analysis, we used the pre-trained classifiers to predict both attentional states for each neurofeedback trial. The raw classifier outputs are shown for an example trial. **b.)** Raw classifier outputs were translated into feedback using the method employed during the actual experiment. Feedback sequences for both classifiers are shown for the same example trial as in **a**. **c.)** Next, we computed the Euclidean difference between the space and feature-based neurofeedback signals. We also computed 1000 permuted differences for unmatched trials. The panel shows the true timepoint-by-timepoint difference for the example trial, together with an example permuted difference time-course. **d.)** The mean Euclidean difference between the space and feature-based feedback signals on matched trials was compared with the permuted distribution of differences calculated for unmatched trials. A p-value was computed for each participant in each session using this method to bootstrap a distribution of dissimilar feedback time courses. **e.)** The resulting p-values are shown for each neurofeedback day (Day1 – 3) and neurofeedback group (Space, Feature). Only 2 out of 213 points could be classified as more similar than would be expected by chance.