Reliability of single subject fMRI in the context of presurgical planning

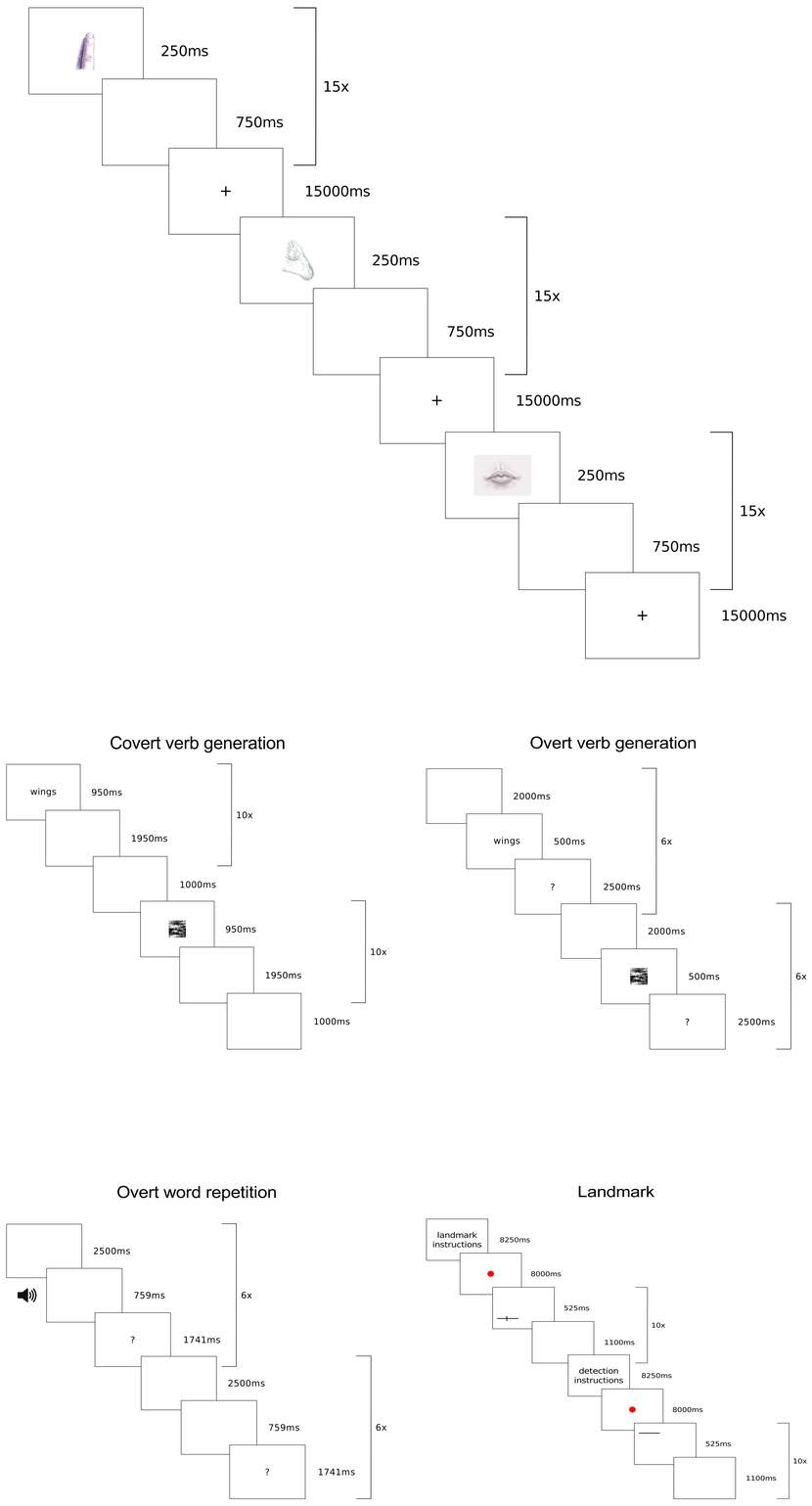
K. Gorgolewski[1], A. Storkey[1], M. Bastin[2], C. Pernet[2]

[1]School of Informatics, University of Edinburgh, UK; [2]Brain Research Imaging Centre, University of Edinburgh, UK;

Introduction

Test-retest reliability of fMRI paradigms has been assessed in many studies [1], however only few have looked at the problem from the context of presurgical planning [2]. fMRI has been used to map the cortex of patients with intracranial tumour in order to improve surgical planning and safety (minimizing chances of neurological deficits). Most neurosurgeons use thresholded maps to identify regions of activation. Therefore we have focused our study on reproducibility of a thresholded single subject t map acquired twice three days apart in normal volunteers undergoing paradigms designed to identify eloquent cortical regions around tumours.

Tasks



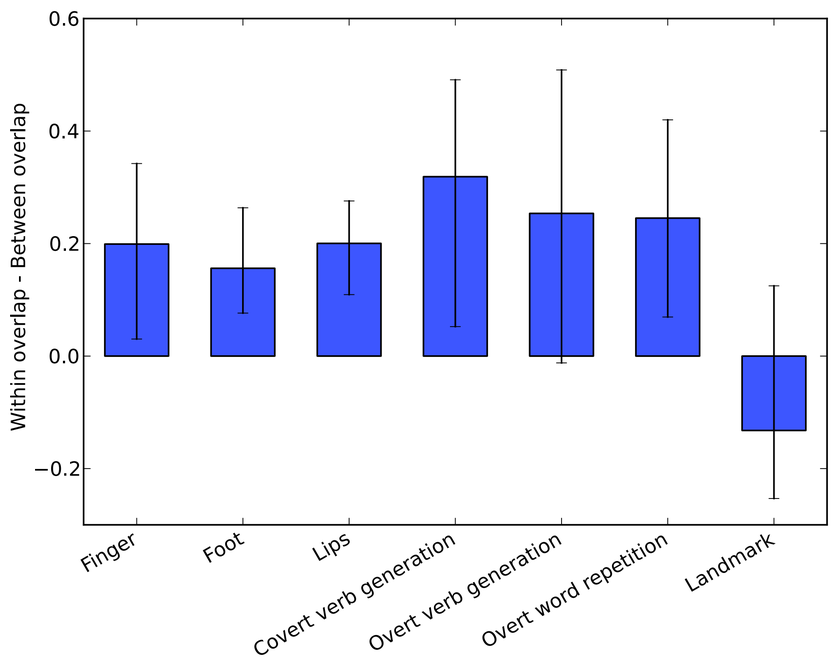
Methods

For each task all acquired images were coregistered and normalized. Functional volumes underwent motion realignment, slice time correction, artefact detection and General Linear Model fitting with motion and artefact regressors. T maps were thresholded using Topological False Discovery Rate (FDR) (q=0.05) with Gamma-Gaussian Mixture Model cluster-forming threshold [3]. Regions of Interest (ROIs) were defined for every task using an atlas (based on knowledge of expected activation from previous studies). For every task Dice's overlap within ROIs was calculated between the two sessions and between each pair of subjects for each session. Two sample bootstrap (n=3000) test of Harrel-Davis median was calculated for each task. Results were corrected for multiple comparisons using Benjamin-Hochberg FDR controlling procedure [4]. Additionally the validity of each task has been confirmed with a second level analysis using Holmes-Friston approach [5].

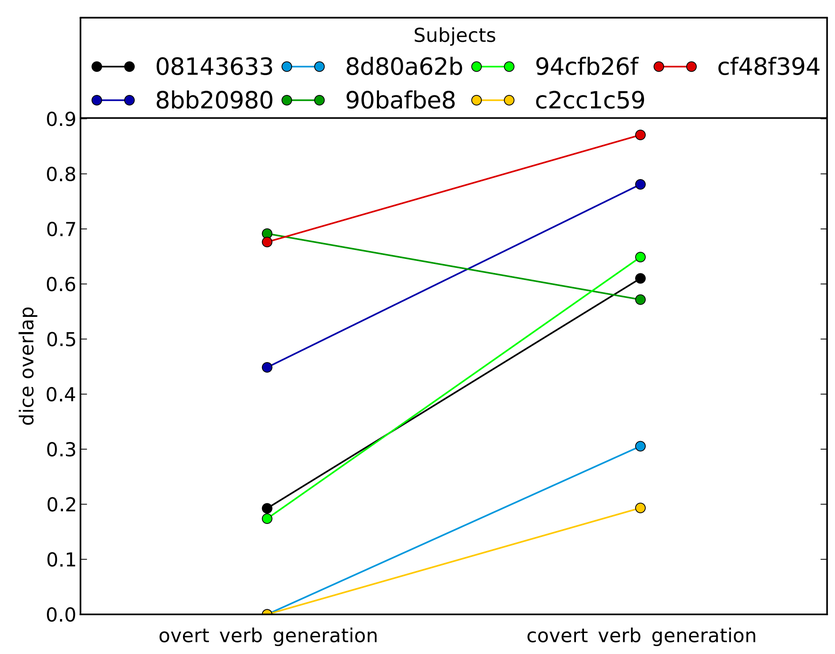
Results

Results cont.

Group analysis results produced activations in areas previously reported in the literature for all tasks. For all motor task contrast (finger, foot, lips) within-subject overlap was statistically significantly higher than between-subject overlap (see Figure 1). Among the language tasks, this was true only for the overt word repetition and covert verb generation. Overt verb generation did not show significantly higher within-subject overlap than between-subject overlap. Performance of the landmark task was the worse out of all the tested tasks. There was no statistically significant difference between the within- and between-subject overlaps. Median within overlap was in fact lower than between overlap.



We have additionally compared within-subject overlaps of the two verb generation task variants paired for each subject. Overlaps were statistically significantly higher for the covert version:



Conclusion

Despite the fact that all of the tasks showed the correct activation pattern in second level analysis, only the motor task, covert verb generation and overt word repetition tests met our reliability criterion. Additionally we have shown that the covert version of the verb generation task is more reliable than its overt counterpart. We have also found that the landmark task is not reliable enough to be used in single subject scenarios and is therefore not useful in preventing post operative spatial hemineglect.

References

1. Bennett, C. M., & Miller, M. B. (2010). How reliable are the results from functional magnetic resonance imaging? Annals of the New York Academy of Sciences, 1191(1), 133-55. 2. McKinsey, R. D., Heredia, G., Fain, S., Meyerand, B., & Tome, W. (2008). Evaluation of Reproducibility of fMRI Maps in Patients with Proven Low-Grade Brain Neoplasms. Proceedings 186h Scientific Meeting, International Society for Magnetic Resonance in Medicine, 2824. 3. Gorgolewski, K., Storkey, A., Bastin, M., & Pernet, C. (2011). Using a Combination of a Mixture Model and Topological FDR in the Context of Presurgical Planning. 17th Annual Meeting of the Organization for Human Brain Mapping.4. Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society. Series B, 57(1), 289-300. 5. Holmes, A., & Friston, K. J. (1998). Generalisability, random effects & population inference. Neuroimage, S754.