Reviewer's report

Title: Medial temporal lobe function during emotional memory in early Alzheimer's disease, mild cognitive impairment and healthy ageing: an fMRI study

Version: 2 Date: 21 August 2012

Reviewer: Mehul Trivedi

Reviewer's report:

Minor Compulsory Revisions:

1) How was MCI classified? Were all MCI patients, amnestic alone or were multi-domain and non-amnestic MCI individuals included? Also, was a 1SD or 1.5SD threshold used to define memory impairment in MCI? This is important because different studies (some cited by the authors) have found that fMRI activation changes in MCI are dependent on the level of functional impairment. Please provide this information.

2) Group differences in NART suggest that the AD group had different premorbid intellectual ability, it would also be useful to provide mean education in Table 1. While this is not a big issue (since the authors included NART scores as a covariate), it will help frame the results of this study in the larger literature on fMRI in AD and MCI.

3) Please insert the correlation values for the analyses that included only the MCI and AD groups (top of pg. 21).

4) When using a cluster level correction threshold, it is common to include the minimum cluster size (k) specified, but the authors did not include this information. Please provide this information.

Discretionary Revisions:

5) Why was no subsequent memory analysis (hits vs. misses) done with the imaging data (see Buckner et al., 2001 - PMID: 11371316; Trivedi et al., 2008 - PMID: 18663302)? Given that the authors have available to them outside the scanner recognition performance and inside the scanner fMRI activation during encoding, a subsequent memory analysis would be extremely informative. I understand that there are two different types of stimuli (neutral and positive), but the authors have combined them in this manuscript, so it could still be done. Conversely, if the authors do not intend to do a subsequent memory analysis, they should mention this in the manuscript and indicate why it was not included.

6) A bigger concern for this reviewer is the statistical thresholds employed in this study. The authors used a voxel level threshold of $p = 0.001$ with an uncorrected (i.e., not FWE or FDR corrected) cluster level correction of $p = 0.05$ using SPM5.
As mentioned above, the authors should include the minimum cluster size associated with the $p = 0.05$. However, a larger question is that there are likely a significant number of false positive voxels in the data set depending on how many voxels are in the data (i.e., using this threshold when there are 10K voxels vs 60K voxels would yield substantial differences in the number of false positives). SPM8 has built in cluster level inferential statistics into the program using FDR and FWE correction methods and this would be more suitable. However, this reviewer has been guilty in the past of using a similar approach as the authors have used in the present study. While I don’t expect the authors to reprocess their data in SPM8, they should acknowledge in the limitations section that there are likely several false positive voxels in their within group and between group activation maps, especially given the small sample sizes. Also, see Bennett et al., 2009 (PMID: 20042432) and Lieberman and Cunningham (2009) (PMID: 20035017) for a worthwhile discussion of statistical thresholds in neuroimaging research.

**Level of interest:** An article of outstanding merit and interest in its field

**Quality of written English:** Acceptable

**Statistical review:** No, the manuscript does not need to be seen by a statistician.