Dartmouth College
Department of Psychological and Brain Sciences
HB 6207 Moore Hall
Hanover, NH 03755

Nature Publishing Group
4 Crinan Street
London N1 9XW, United Kingdom

October 12, 2018

To the Editors of *Nature Neuroscience*:

In August of 2017 we submitted a manuscript for consideration as a research article in *Nature Neuroscience*, entitled *Towards Human Super EEG* (NN-A60453-T). Our paper presents a method for estimating neural activity throughout the human brain at high spatiotemporal resolutions, using recordings from a relatively small number of implanted electrodes. While our original submission was rejected during the review process, we wish to appeal that decision in the hopes that you might consider a substantially revised and updated version of our paper.

In brief, while the original reviewers found the work to be of potential interest, they raised concerns about the strength of the conclusions, some aspects of our methods, and the robustness of the data. We have now done a substantial overhaul of our paper that we believe addresses the reviewers' concerns, including the following major changes:

- We have improved several aspects of our methods and analyses, which has cleaned up the prior results.
- We have replicated our results (originally reported on a single large ECoG dataset) using a second large ECoG dataset, thereby demonstrating the robustness of our findings and data. This second dataset also allowed us to directly test several of the speculative claims we had raised in our prior submission.
- We have improved our statistical tests (largely following the original reviewers' suggestions), demonstrating our main findings more clearly and cleanly.
- We have addressed concerns about a PCA-based analysis that we had included in our prior submission (Reviewers 1 and 2 had raised concerns about the strength of that analysis, and we agree with the reviewers' concerns). That analysis was tangential to our main set of results, and so we have chosen to remove that analysis entirely rather than revamp or revise it.
- We have released an open-source Python toolbox that implements our main method, in addition to the code we used to carry out our analyses, which we hope will lend an additional measure of transparency to the review process. (Note that we have added a third author to our manuscript, reflecting their contribution to these software development efforts.)

Our revised manuscript makes several contributions to the literature. First, as in our original submission, the approach we present examines the extent to which ECoG recordings from one set of brain locations may be used to infer activity patterns throughout the rest of the brain. We believe our approach could be transformative for obtaining high spatiotemporal resolution activity patterns throughout the brain (i.e., at spatial resolutions comparable to fMRI and

temporal resolutions comparable to EEG and ECoG). We test our approach by holding out portions of the ECoG data and comparing that held-out data to the predictions about activity patterns in those regions generated by our model. We show that, by training a model on data from a large number of patients, one can accurately "fill in" held-out data at arbitrary locations, *even when those locations are distant from where the training-set electrodes were implanted in that patient's brain*. By replicating this result across two datasets, the claims in our original submission have been strengthened.

We also present two important new findings, which help to address ongoing questions in the brain networks literature. Specifically: to what extent is the correlational structure of our brain activity (a) common across people and (b) common across tasks? In one set of analyses, we show that training our model using data from different patients substantially improves its predictive performance, *even when compared with models trained directly on the subject whose brain patterns we are predicting*. This is a surprising result that speaks to the similarities in different people's functional connectomes. In another set of analyses, we systematically train and test our model on data from two different experimental tasks. We show that, although our model achieves good across-task performance (indicating that some aspects of human functional connectomes are similar across tasks), we observe the best performance when our model is trained and tested using data from the same task. This result is consistent with a growing literature on task-specific brain network dynamics. However, whereas prior work on across-task neural predictions has attempted to estimate *average* activity and functional connectivity patterns, our approach estimates the full spatiotemporal dynamics of held-out data, which may then be used to examine single-trail or single-moment neural patterns, from individual patients.

We have included point-by-point reviewer responses below (the reviewers' comments are in *italics* and our responses are in **bold**). Thank you for considering our appeal.

Sincerely,

Jeremy R. Manning
Email: jeremy.r.manning@dartmouth.edu
Phone: 603-646-2777

*Reviewer #1:*

*Assessment*
*The present study offers a method to provide increased spatial coverage to high spatial and temporal resolution data (i.e., ECoG data). While the method appears technically sound and clever, the validation of the method seems somewhat biased (see below comments). Further, the reported results make me skeptical of the exact contribution of the method (e.g., what is the benefit of having the observed time series of a held-out location in a subject if one already has the group covariance matrix?).*

**Whereas the group covariance matrix provides information only about how signals covary on average, the full time series of signals at held-out locations provides information about the moment-by-moment *dynamics*. For example, the correlational structure of the data on short timescales may look substantially different than the average group covariance matrix.**

*Given that the inferred/predicted activation patterns are most similar to the observed activation patterns "... for patients whose brain data are the most redundant..." (pg. 14; Fig. 6), inferring redundant activity patterns seems trivial.*

**The analysis the reviewer is referring to was presented poorly and does not add to our main story; we have removed it from our revised manuscript. In brief, we were attempting to estimate the effective "rank" of each patient's data, and show that patients with lower rank data were better predicted by our model. However, there are a number of problems with our prior approach to answering this question-- for example, patients with fewer electrodes are necessarily "lower rank" than patients with many electrodes. We have instead added a set of supplemental analyses examining which properties of the data predict reconstruction accuracy (Fig. S4).**

*This potential triviality is compounded by the fact that the statistical model (as formalized in Equation 6) predicts the time series of a held-out electrode using the data from an electrode from the same temporal window (see Major comment 1 for more details).*

**See response to major comment 1 below.**

*While I find the framework and approach of potential interest, I am not convinced that the method in its current form provides an advance to our current understanding of whole-brain dynamics. Previous studies and models have been able to predict task-evoked activation patterns on held-out subjects (Tavor et al. 2016) and/or held-out runs using only resting-state functional connectivity in fMRI (Cole et al. 2016). The current manuscript does neither, and only predicts on a held-out location during the same time window.*

**We thank the reviewer for directing us to these papers, and we have added citations of these studies to our revised discussion section (page 19). In brief, those prior studies are both concerned with predicting *static average* task-evoked activity and functional connectivity patterns using out-of-task data. Our approach predicts *dynamic* activity patterns from held-out brain locations. In other words, the above studies predict task-averaged snapshots for held-out tasks (where different tasks are performed at**

**different timepoints), whereas our work is aimed at predicting held out spatial patterns that unfold over extended time intervals.**

*While the method may produce the normalized voltage time series of that run, I'm skeptical due to the potential circular nature of the analysis (detailed below).*

**See response to major comment 1 below.**

*Further, the proposed name of the method – super EEG – is greatly misleading, as the inference provided from the method is very far from the intended use of the term "super EEG". It would be necessary to completely remove the term "super EEG" (or perhaps mention it only briefly in passing, as a metaphor) for this method and manuscript to be accepted by the scientific community.*

***SuperEEG* is a made up term from the science fiction literature that has no standard meaning in the neuroimaging literature. We titled our original paper "*Towards* Human SuperEEG" to reflect that our approach makes an incremental step towards the science fiction goal (now referenced in a footnote on page 3 of the Introduction) of inferring full-brain activity patterns at arbitrarily high spatiotemporal resolutions. In fact, our proposed method *does* (mathematically) generate predictions of full-brain activity patterns at arbitrarily high spatiotemporal resolutions, thereby keeping with the "spirit" of the SuperEEG name. What we lack in the current manuscript is an ability to directly test those predictions at arbitrarily high resolutions (since we are limited by current recording technologies). Further, we note in our Discussion (page 20) that we hypothesize that the predictions generated by our model will break down at the level of single neuron action potentials.**

**For the reasons above, we disagree that the term SuperEEG is misleading *per se*, and so we have elected not to remove the term entirely from the manuscript. However, we concede that our emphasis on the term could be interpreted as overbearing, and we have therefore toned down our use of the term SuperEEG substantially throughout the entire revised manuscript.**

*Major:*
*1. The demonstration of the analysis is potentially circular. Specifically, if two regions are spatially adjacent to each other, the time series of the to-be-inferred region can contaminate the covariance matrix due to volume conduction/spatial autocorrelation (i.e., variable Ya in equation 6 for any location alpha close to location beta when predicting Yb). (Figure 1D seems to indicate that spatially adjacent locations are highly correlated with each other.) Thus, if a signal from the to-be-predicted region was not carefully removed from an observed signal, the approach would be made trivial due to circularity. Adding to the likelihood that the analyses are circular, zero-lag correlations were used. It is well known that volume conduction artifacts have essentially zero lag due to signals propagating at the speed of light – much faster than real neural signals. It would be much more convincing that the analyses are not circular if non-zero lag functional connectivity measures were used [see for example Barnett L, Seth AK (2014) "The MVGC multivariate Granger causality toolbox: A new approach to Granger-causal inference". Journal of Neuroscience Methods. 223:50–68.http://doi.org/10.1016/j.jneumeth.2013.10.018 and Nolte G, Ziehe A, Nikulin VV,*

*Schlögl A, Krämer N, Brismar T, Müller K-R (2008) "Robustly Estimating the Flow Direction of Information in Complex Physical Systems". Phys Rev Lett. 100:4.http://doi.org/10.1103/PhysRevLett.100.234101].*

**The reviewer is correct that activity at spatially adjacent locations is often correlated (in fact, that is one of the assumptions of our model). We have added a strong test of the reviewer's concern. Specifically, we asked: suppose most of the predictive power of our model comes from the sorts of volume conductance artifacts the reviewer is describing, rather than by the learned full-brain correlation structure, including between distant structures, that we are attempting to model? To test this, we compared the model's predictions under two conditions. In a first condition (*across-patient*), the model learned the full-brain correlation structure using only data from *other* patients (i.e., excluding a single held-out patient whose data we were reconstructing). In a second condition (*within-patient*), we trained the model using only data from the *same* patient whose data we were reconstructing (i.e., we held out data from a single electrode, and trained the model using the remaining data from that patient). If the model's predictions were driven primarily by volume conductance artifacts, or even if there were substantial across-patient differences in functional correlations, then the within-patient condition should have yielded better reconstruction performance. However, we in fact observed that the across-patient predictions were far more accurate (we replicated this result across our original dataset, and also in a new dataset that we added to our revised manuscript-- Figs. 3, S1, S2, and S3).**

**The reviewer's second point, about zero-lag correlations, is also well-taken. We agree that our general approach could be extended to incorporate lagged correlations, or more sophisticated measures such as Granger causality. However, we disagree that our focus on zero-lag correlations implies circularity. First, we have avoided making any claims of "causality," and we do not view the correlation models as (necessarily) reflecting information transfer or communication in the traditional sense. Rather, we are leveraging observed correlations as a convenient source for making predictions. These correlations could reflect direct connections (albeit imperfectly estimated, due to our not having accounted for signal lags in our model), or indirect connections mediated by "third party" structures (which might appear as near "zero lag" correlations, e.g. if the third party structure had similar lags to each of the structures under consideration).**

**We suspect that additional signal might be extracted if we were to incorporate non-zero-lag correlations (e.g., by accounting for white matter path lengths between different structures). However, we were able to achieve good accuracy even without accounting for signal propagation delays (perhaps because we used recordings with high sampling rates, and because the signals themselves exhibit strong autocorrelations-- so the impact of being "off" by a small lag is lessened). For the sake of keeping our model and story simpler, we have therefore opted not to model signal propagation delays in this manuscript.**

*2. It would be helpful for ruling out circularity to test if the method is still effective when electrodes near (e.g., within 1 cm) each to-be-predicted electrode are excluded from the analysis. Ideally this would be based on some principled estimate of the size of volume conduction artifacts.*

**The "within-patient" versus "across-patient" test described above provides an even stronger test of this concern. In other words: when we exclude *all* data from the to-be-predicted patient's electrodes when training the correlation model, we are nonetheless able to reconstruct their held-out activity patterns.**

*3. Related to the previous comment, the choice of lambda in the radial basis function seems important. Unlike how lambda was described, however, a large lambda might be important for reducing the chance that volume conduction artifacts are driving the prediction results. It would thus be important to ensure that radial basis functions do not overlap between each to-be-predicted electrode and electrodes used to predict that electrode. Is this the case?*

**Again, our "within-patient" versus "across-patient" test described above provides an even stronger test of this concern. While a wide RBF kernel might lead to "leakage" between signals recorded from a patient's other electrodes, this cannot happen across patients.**

*4. It is stated that "We compared this distribution of correlation coefficients to the distribution of across-session average correlation coefficients between the recordings from each electrode and its nearest neighbor from the same patient (paired t-test between z-transformed correlation coefficients: t(4148) = 2.36, p = 0.018)." This seems like an especially important test. Thus, more information about it should be provided. What are the values that contributed to the t-test? Can they be summarized somehow, such as mean and standard deviation? Is there any possibility of a random effects test, with the degrees of freedom based on subjects rather than electrodes?*

**We have improved this test by focusing on comparisons between the within-patient versus across-patient models described above. We now report the mean reconstruction accuracy obtained from each model, histograms showing the full distribution of reconstruction accuracies across all electrodes (Figs. 3, S1, S2, S3), and across-patient t-tests comparing the average reconstruction accuracies for each patient obtained using different training sets (pages 12--15).**

*5. It is unclear how much more variance the predicted time points from the model provides over and above the existing, observed data for a given subject. For example, Figure 2B provides a 2D histogram of the observed versus reconstructed voltages for every electrode in the sample. However, the p-value reported here assumes that the mean of the null distribution is 0 (where the reported statistics are r = 0.95; p < 10-10). However, given the findings in Figure 6 and the corresponding inference that 'brain activity is highly redundant', it seems naïve to assume that the null distribution of the correlations in Figure 2B would be 0. Since electrodes contain high amounts of shared variance (particularly between regions close to each other; Figure 1D), and the method uses data from the same run to infer data of a held-out electrode in the same run, it seems that using a correlation of r = 0 as the null hypothesis seems like an incorrect assumption. I would suggest performing permutation tests to estimate the underlying null distribution and to test the significance of the r-values in Figure 2B. For instance, it would important to scramble/permute the correlation matrix many times to establish that the prediction accuracies are significantly lower than the reported results.*

**We agree, and we have therefore revamped our framework for testing our model. In our revised manuscript, we use the within-patient model as a "benchmark" to help elucidate how much predictive power is gained by integrating data across patients (pages 11--12).**

*6. A strong assumption of the approach is the statistical model, which is based on Gaussian process regression. This assumes that the underlying data are normally distributed, including the inferred time series of the held-out electrodes. While the authors z-normalize their data prior to fitting the model and generating predictions, it is well known that electrophysiological signals are non-Gaussian (Buzsáki and Mizuseki 2014). The authors should report the underlying distribution of the electrophysiological data prior to z-normalizing. The authors should also comment on this discrepancy in the discussion. It may be beneficial to also report/compare the current statistics with a (non-parametric) rank correlation (in Figures 2, 3, and 5).*

**The reviewer's point about an appropriate "null distribution" is well-taken. As noted in our response to the previous comment, we now use models trained within-patient as a benchmark for comparing our held-out data results (Fig. 3, pages 11--12).**

*7. The overall tone of the manuscript overstates the results. Besides needing the remove the term "super EEG", there are several places in the manuscript where terms like "exciting" are used. These instances can reduce the reader's enthusiasm for the results since they suggest the authors did not at least attempt to achieve objectivity in their analyses. It would be helpful to let the data and results speak for themselves, rather than potentially exaggerating the importance of the results with flashy terms and unnecessarily biased language.*

**We appreciate this concern. We have toned down our language throughout our revised manuscript to better let the data and results "speak for themselves."**

*Minor:*
*1. It is stated (pg. 5): "We use spatial correlations in the data to estimate these anatomical connections." It seems that these are rather temporal correlations between spatial locations (i.e., functional connectivity estimates). Is this correct?*

**Yes, good catch. Fixed (page 7).**

*2. Correlation values are used to weight the predictions, similar to previous approaches (which should be cited): [Tavor I, Parker Jones O, Mars RB, Smith SM, Behrens TE, Jbabdi S (2016) "Task-free MRI predicts individual differences in brain activity during task performance.". Science. 352:216–220. PMID: 27124457 http://doi.org/10.1126/science.aad8127 and Cole MW, Ito T, Bassett DS, Schultz DH (2016) "Activity flow over resting-state networks shapes cognitive task activations.". Nat Neurosci. PMID: 27723746 http://doi.org/10.1038/nn.4406]. Importantly, both of those studies showed that (unnormalized) multiple regression estimates were far superior for making predictions than correlation values. It seems that a multivariate autoregressive model would be more effective in the case of ECoG data (with the temporal lag helping rule out conduction artifacts). It might be useful to use such a statistical model.*

**This is an interesting idea for follow-up work-- thanks for the pointer! We have added citations of these studies (and we discuss them on page 19).**

*3. To estimate the contribution of different brain regions to a single electrode (via volume conduction), the authors use a radial basis function at the location of each electrode to estimate how structures around the electrode contribute to the recording. However, this seems to have many assumptions that are not carefully addressed in the present manuscript. The authors reference some of their previous work (i.e., ref. 13), but the implications of their choice of lambda = 20 should be explicated in the current manuscript. (It seems somewhat unlikely that all electrodes contain the same amount of spread from surrounding locations.) These assumptions are of particular importance due to the fact that the authors are trying to predict the time series of a held-out electrode using all other electrodes during the same run. Thus, any spread of signal from the same source region (especially between adjacent regions due to spatial autocorrelation) could confound their predictions.*

**We have added additional details about how we selected a value of lambda = 20 (page 6--7). In our prior work (Manning et al. 2014, 2018) we developed a model that (essentially) decomposes fMRI data into sums of RBFs. In those papers, we applied our model to several fMRI datasets and examined the distribution of best-fitting RBF widths. Using the notation in our current manuscript, the median RBF widths tended to be around the equivalent of 20 MNI units. We also applied another general heuristic to Dataset 1 to further tune the lambda parameter. In particular, we constructed a set of RBF functions centered on each Dataset 1 electrode, and then evaluated the sum of those functions at the location of each voxel in the 4 $mm^3$ MNI (standard) brain. When we set lambda = 20, this seemed to provide (roughly) the smallest RBF width that also covered nearly the entire brain volume with non-zero weights. (We then used the same value of lambda for the remaining analyses in our paper, including analyses of Dataset 2.)**

**Regarding the reviewer's concern that spreading signal from surrounding electrodes could confound our predictions: we used exactly this model (called a "within-patient" in our revised manuscript) as our benchmark to comparatively evaluate the performance of a model that incorporates correlational information learned from other patients (i.e., an "across-patient" model). We found that, in every test we ran for both Datasets, the across-patient model substantially out-performed the within-patient model.**

*4. The choice of time points (b = 25000) for estimating the covariance matrix for a given subject during a given run seems arbitrary (aside from the mention that any number > 25000 would be computationally intractable). Have the authors considered down-sampling time points (e.g., temporal smoothing) to obtain a covariance matrix sampled over a longer time duration? Previous simulation work in complex systems have shown that different community structures emerge from estimating the correlation/covariances of a network over different time scales (Arenas et al. 2006). Thus it seems prudent to at least establish the implications of estimating the covariance matrix over a constant time window. (For example, is it possible that estimating the covariance matrix over smaller windows provides better predictions of the voltage dynamics at shorter time scales?)*

**We explained this point poorly in our previous draft. We have added a note (page 10) that our model is timepoint invariant, and therefore to make the computations more tractable**

**one can apply out reconstruction algorithm (Equation 12) independently for every timepoint (or blocks of timepoints), and still obtain numerically equal predictions to estimating the full timeseries in a single block.**

*5. As a general comment, providing a code demo/tutorial may be helpful, especially in the case of a new method.*

**Shortly after we submitted our previous manuscript, we released an open-source Python toolbox that implements our full approach and allows users to use SuperEEG to reconstruct full-brain activity patterns using any ECoG dataset (link). We have also released a series of SuperEEG tutorials (tutorial 1, tutorial series 2), a conceptual video (link), a tutorial video (link), as well as the analysis code used to generate all of the figures in our paper (link). Unfortunately, we neglected to include any of these links to this material in our original submission; we have now corrected that omission (page 21).**

---

*Reviewer #2:*

*Remarks to the Author:*
*The authors present a method for reconstructing the "predicted" intracranial EEG/ECOG timecourse of a signal at an electrode location not actually measured. They do this by utilizing the correlation structure between electrodes measured in other patients (as an estimate of connectivity) and volume conduction using point-spread functions. They show that this algorithm can predict activity on left-out electrodes with remarkable accuracy when averaged across many hours of data.*

*While interesting, this methodological paper is preliminary and includes neither the depth of technical advance nor biological insight necessary to warrant publication in this venue. While certainly an interesting method, it is unclear what aspects of brain activity this method predicts and whether such 'predicted' activity would be useful.*

**See response to Major issues 1 below.**

*A comment on the title: a super resolution technique refers to actual measurements made, not to predictions derived from other measurements. Thus this is not a step towards "super EEG" like that done in super resolution microscopy or similar.*

**With respect to super resolution techniques being defined (or implied) as being limited to "actual measurements," we respectfully disagree with the reviewer. Although some super resolution techniques do rely on actual measurements as the reviewer suggests, this is not true of *all* super resolution techniques. For example, in our own prior work (e.g., Manning et al. 2009, *Visual Neuroscience*) and older work on optical super resolution (e.g. Harris et al., 1964, *Journal of the Optical Society of America*) Bayesian estimators are applied to comparatively low resolution sensor measurements to obtain higher resolution estimates, even at locations where no sensors were placed. Our current method bears some mathematical resemblance to the optimal estimators derived in that prior work.**

**The above being said, taken with Reviewer 1's additional concern about the term SuperEEG, we have significantly toned down our use of the term throughout our revised manuscript, including changing the title of our manuscript and substantially editing the abstract and introduction.**

*Major issues:*
*1.What types of activity can this technique predict? What is evaluated are long stretches of data (many hours). But the correlations in such activity are going to be dominated by high-amplitude low-frequency oscillations, which are coherent across large parts of the brain. But this does not mean that this technique would be able to predict specific task-evoked activity, which is substantially more local and often carried by higher frequencies. Without this, it is unclear if these correlations are driven only by global low-frequency activity. To be convincing, what is needed is at the very least: 1) analyze task-evoked activity and to show that task-relevant variables are locally encoded in. 2) analysis if different frequency bands and tests of meaningful predictions known to be possible (i.e. beta in the motor system). Simply correlating broad-band activity is going to be entirely dominated by very low frequencies due to 1/f nature of the signal, so this would have to be cancelled out.*

**We have added a series of analyses intended to assess the task specificity of our estimates (pages 13--15, Fig. S3). We found that reconstructed activity patterns were most accurate when our model was trained using within-task data. This indicates that our model is picking up on task-relevant information.**

**We agree that a careful analysis of which frequency bands our model is picking up on would be useful, and in followup work we plan to more fully characterize which aspects of the data SuperEEG explains well. However, such an analysis is beyond the scope of what we can reasonably fit into this manuscript, where our focus is on presenting the method and demonstrating its performance by showing that SuperEEG may be used to recover held-out data. Further, the current set of analyses are very computationally intensive (primarily in terms of space required to store the resulting reconstructions). Breaking down the training and testing by $n$ frequency bands would increase the computational cost by a factor of $n$, which would be non-trivial to accommodate. Nevertheless, we did carry out some preliminary analyses to this effect. For example, we examined whether the data acquisition (sampling) rates for each recording session were correlated with reconstruction accuracy ($r = -0.14$, $p < 10^{-20}$). This negative correlation (though weak in magnitude) implies that our approach may be picking up more on lower frequency aspects of the signal. And as the reviewer notes, correlating raw signals as we have done in our manuscript is likely to be dominated by the (much larger amplitude) low-frequency components of the recordings.**

*2. No insight is provided into what elements of the correlation matrices drive the ability to predict. For example, is prediction only high if a measured electrode is either nearby the one that we would like to predict or at a similar location in the other hemisphere? The analysis is much too coarse to provide any biological insight. The brain areas are not labeled for the connectivity matrices, so I cannot assess whether they look anything like expected from other methods.*

**We have added several analyses to give insights into these questions. First, to provide some additional insights into discovered structure in the correlation matrices, in Fig. 1 we performed a *k*-means clustering analysis and colored each electrode according to its cluster label (we used *k* = 7, following Yeo et al. 2011, *Journal of Neurophysiology*). We also carried out a series of analyses examining the impact of electrode density (i.e., the number of nearby electrodes) on reconstruction performance (Fig. 4, pages 15--16). Finally, we carried out a series of analyses examining which specific brain areas were most informative about ongoing activity in *other* brain areas (Fig. 5 and S5, pages 17--18).**

*3. PCA analysis. The fact that PCA allows a low-dimensional projection does not mean that brain activity is low-dimensional. This is because this analysis is dominated by high-power but low-frequency activity. Detailed fine-grained analysis of different brain structures has recently revealed remarkably high dimensionality at the single-neuron level, and this depends on the brain area being studied (i.e. M1, PPC, pre-frontal; work by Churchland et al, Shenoy et al, Fusi et al).*

**[COPIED FROM RESPONSE 4 TO R1] This analysis was presented poorly and does not add to our main story; we have removed it from our revised manuscript. In brief, we were attempting to estimate the effective "rank" of each patient's data, and show that patients with lower rank data were better predicted by our model. However, there are a number of problems with our prior approach to answering this question-- for example, patients with fewer electrodes are necessarily "lower rank" than patients with many electrodes. We have instead added a set of supplemental analyses examining which properties of the data predict reconstruction accuracy (Fig. S4).**

*4. No detailed methods on how the data was acquired, re-referenced, grounded etc so I'm unable to judge.*

**We analyzed two previously collected ECoG datasets. We provide a full list of studies reporting the detailed data acquisition methods for both datasets (pages 3 -- 4). In brief:**

> **From Manning et al. (2011), reporting on data collection methods for Dataset 1:**
>
> ***ECoG signals were recorded referentially using a Telefactor, Bio-Logic, XLTek, Neurofile, or Nicolet electroencephalographic digital video-EEG system. Depending on the amplifier, signals were sampled at 200, 256, 500, 512, or 1,024 Hz.***
>
> **The data were referenced to a common contact placed either intracranially, on the scalp, or on the mastoid process.**
>
> **From Ezzyat et al (2017), reporting on data collection methods for Dataset 2:**
>
> ***Intracranial data were recorded using one of the following clinical electroencephalogram (EEG) systems (depending the site of data collection): Nihon Kohden EEG-1200, Natus XLTek EMU 128 or Grass Aura-LTM64. Depending on the amplifier and the preference of the clinical team, the signals were sampled***

***at either 500, 1000 or 1600 Hz and were referenced to a common contact placed either intracranially, on the scalp or mastoid process.***

---

*Reviewer #3:*

*Remarks to the Author:*
*Owen and Manning describe an approach for combining human intracranial EEG (iEEG) signals across subjects to calculate estimates of responses in unrecorded locations for a given subject. The motivation is driven by the inherent brain coverage limitations imposed by iEEG, which has otherwise high spatial and temporal resolution of underlying local field potentials. Recordings of iEEG are motivated by their clinical application and are thus typically focused on limited brain regions; having access to the precision of iEEG across a wider expanse of the brain would be useful for both clinical and non-clinical work, keeping in mind the inherent restrictions and limitations that accompany recordings in the patients. The authors bring together multiple datasets of iEEG patients, each collected during verbal list-learning (study and recall), wherein iEEG electrode placement is sampled differently across the patient's brains. Using a Gaussian process regression method applied to the data-set the authors are able to predict responses in held out electrodes, thus suggesting that the approach can be used to predict response profiles in electrode-absent locations for an individual with limited electrode coverage.*

*iEEG datasets are becoming more prevalent and efforts to provide tools and methods to more broadly utilize the information are certainly welcome. Despite my enthusiasm for the overall motivation, a number of features of the presented research leave me reluctant to accept the general take home message and generalizability of the approach. A more in depth exploration of the data and approach are required in order to understand its potential utility and its limitations. In addition, the approach relies on a number of assumptions regarding cortical organization and subject variability that I believe require consideration. I make comments towards these impressions below.*

*(1) A large assumption made throughout the report hinges on the consistency of cortical organization across individuals (e.g., area parcellation). Many recent reports focusing on cortical organization in individuals have reinforced reasons to believe otherwise, by demonstrating substantial variability in the localization of functionally areas/regions across individuals (e.g., Tavor et al, 2016 for differences using task and resting-state BOLD; Caspers et al., 2006 for differences revealed using measurements of cytoarchitectonics). This variability is present in even 'homogenous' cohorts; I would expect it to be exacerbated in the group of patients studied here who likely vary in etiology/impairment, age, medication status etc.*

**We carried our a series of analyses examining whether functional correlations were sufficiently similar across individuals to provide meaningful predictive power (Fig. 3, S1, S2, S3, pages 12--13). Although the reviewer's points about across-subject variability are well-taken, our results indicate that it is nonetheless beneficial for our model to learn about the correlational structure of full-brain activity patterns by incorporating data from different patients (and, to an extent, different tasks). We have also added a discussion of**

**these points, citing the Tavor et al. 2016 study and several other related studies (pages 18--19).**

*One relevant aspect of the present results reveals the (somewhat expected) observation that the locations with densest sampling exhibit the best reconstruction, presumably because they have the most available information (e.g., the hippocampus and surrounding medial temporal lobe). At present It's not clear to what extent functional variability relates to the sampling variability and density; are the locations that are less well sampled also those that exhibit more variability in parcellation across participants? I realize this is a challenging question, but without deeper insight towards resolving this issue it is difficult to understand the nature of the information that is extracted, and measures of confidence across locations for a given individual, which I think are central aspects of the paper.*

**We have added several analyses examining these questions (Figs. 4, 5, and S5, pages 15--17).**

**First, we improved our analysis relating sampling density to reconstruction accuracy and replicated the results in a second ECoG dataset (Fig. 4). There were a number of problems with our prior characterization of electrode sampling density. When we simplified our density calculation to reflect the proportions of electrodes contained in a sphere centered on each voxel, we obtained much more cleanly structured maps that were common across both datasets we examined. We also found a (weak) *negative* correlation between sampling density and reconstruction accuracy, providing some evidence against the reviewer's proposed MTL explanation.**

**We also carried out an analysis intended to provide insights into which electrode locations provide the most information about activity patterns at *other* locations (Fig. 5, S5). We found that, across the two datasets we examined, some similar locations were identified by our analysis as being especially informative about full-brain activity patterns. These primarily comprised structures in the dorsal attention network and default mode network that are known to be heavily interconnected with much of the brain's gray matter (pages 17--18).**

*(2) The results and methods are based off analysis of list-learning accompanied signals. I recognize the challenge in synthesizing this type of data. However, at present It's not clear how generalizable the observations would be to other tasks states. The authors discuss this issue in the discussion section but I think in order to have substantial impact greater exploration of variability/consistency across 'states'/sessions would need to be included. Without this, to me, the broader take home converges with previous reports that have revealed consistencies of field potential and BOLD signals across participants in specific locations under a given task condition.*

**We have added several analyses assessing the task-specificity of our results. First, we extended the time intervals we examined from Dataset 1 (previously the only dataset we examined) to include times when the patients were participating in a list-learning task, in addition to times when they were not participating in any explicit task. This allowed us to assess the task-general predictive power of our approach. We also added a series of task-specificity analyses on Dataset 2 (newly added to the manuscript), which included data recorded as each patient carried out each of two different memory tasks. We found**

**that, although our model's predictions were best when applied to same-task data, we also achieved good across-task predictive performance (Fig. S3, pages 13--15). Therefore, despite across-task differences in functional correlations, there nonetheless remain sufficient commonalities in functional correlations across different tasks to support our model's high across-task reconstruction accuracy. We have also added a discussion of these ideas (pages 18--19).**

*(3) The authors note that one constraint in iEEG, in addition to the locations of the electrodes, is the number of electrodes that are implanted. I think it would be important to explore this variable to attempt to understand how number of sampled subject electrodes impact estimates of his or her responses in absent electrodes.*

**We have added a series of supplemental analyses to better understand how different properties of the data affect reconstruction accuracy (Fig. S4). In brief, across the two datasets we examined, we found no consistent relation between the number of implanted electrodes and reconstruction accuracy.**