NeuroImage 197 (2019) 435–438

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/neuroimage

Classification of temporal ICA components for separating global noise from fMRI data: Reply to Power

Matthew F. Glasser [a,b,c,*], Timothy S. Coalson [a], Janine D. Bijsterbosch [d], Samuel J. Harrison [d,e], Michael P. Harms [f], Alan Anticevic [g], David C. Van Essen [a], Stephen M. Smith [d]

[a] Department of Neuroscience, Washington University Medical School, Saint Louis, MI, 63110, USA
[b] Department of Radiology, Washington University Medical School, Saint Louis, MI, 63110, USA
[c] St. Luke's Hospital, Saint Louis, MI, 63017, USA
[d] Centre for Functional MRI of the Brain (FMRIB), Wellcome Centre for Integrative Neuroimaging, Nuffield Department of Clinical Neurosciences, University of Oxford, John Radcliffe Hospital, Headley Way, Oxford, OX3 9DU, UK
[e] Translational Neuromodeling Unit, University of Zurich & ETH Zurich, Wilfriedstrasse 6, 8032, Zurich, Switzerland
[f] Department of Psychiatry, Washington University Medical School, Saint Louis, MO, USA
[g] Department of Psychiatry, Yale University School of Medicine, 300 George Street, New Haven, CT, 06511, USA

A B S T R A C T

We respond to a critique of our temporal Independent Components Analysis (ICA) method for separating global noise from global signal in fMRI data that focuses on the signal versus noise classification of several components. While we agree with several of Power's comments, we provide evidence and analysis to rebut his major criticisms and to reassure readers that temporal ICA remains a powerful and promising denoising approach.

1. Introduction

We thank Power (2019) for his perspective on our recent paper about using temporal Independent Components Analysis (ICA) to separate global, largely respiratory, artifact from fMRI data (Glasser et al., 2018). We start with points of agreement with Power: (1) We agree that global noise is an important problem for fMRI data. For resting-state fMRI, this issue has been widely discussed, including by Power (Power et al., 2017, 2018; Burgess et al., 2016), and we also highlighted its effects on task fMRI data (Glasser et al., 2018). (2) We also agree that temporal ICA (tICA) is a promising approach for removing global noise from fMRI data and provides a quantitative platform upon which differing hypotheses about the nature of global respiratory signals can be explicitly tested (i.e., depending upon how the temporal ICA components are classified). (3) We agree that the optimal component classification might eventually be revised from the one originally presented in the paper after further study of these phenomena. We hope that this dialogue leads to a deeper, neurobiologically-principled experimental and theoretical basis for fMRI data denoising. However, we disagree with a number of Power's comments and found the critique insufficiently clear and lacking balance in several critical respects. We aim to clarify these here, before considering what we believe to be the heart of the debate.

1.1. Preliminary points

Power frames his critique in a way that does not explicitly state that his core concerns are about our chosen classification of several components as 'signal' instead of 'noise'. While Power does discuss the signal versus noise classification of some tICA components without specifically naming them, a reader might reasonably infer from the commentary and its title that Power doubts whether the entire tICA framework can separate global noise from global signal, regardless of the classification chosen, and that he might believe that global signal regression (GSR), an approach that he has previously advocated (Power et al., 2014; Burgess et al., 2016), would outperform tICA. Thus, we first briefly address any uncertainty over these issues and provide more detail in Section 1 of the Supplement to this Reply, where we show the following: i) The features of the task greyplot that Power highlighted (in his Fig. 1) actually represent the blocked task design's neural effect on task positive regions of the cerebral cortex, illustrating the pitfall of assuming that all structure within greyplots that may be temporally coincident with respiratory noise is artifactual. These findings (also illustrated below in Fig. 1) make clear that Power's later arguments based on the expected timecourses of event-related task fMRI neural signal do not generalize to blocked task fMRI neural signal, nor likely to resting-state fMRI. ii) There are numerous examples of greyplot bands

* Corresponding author. Department of Radiology, Washington University Medical School, Saint Louis, MI, 63110, USA.
E-mail address: glasserm@wustl.edu (M.F. Glasser).

https://doi.org/10.1016/j.neuroimage.2019.04.046
Received 27 January 2019; Received in revised form 11 April 2019; Accepted 17 April 2019
Available online 24 April 2019

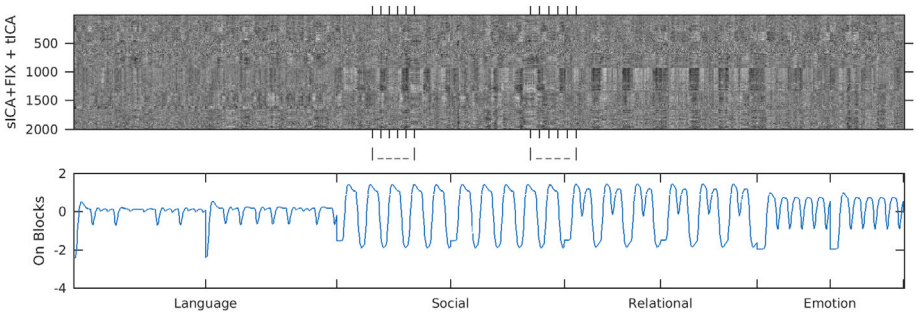

Fig. 1. Temporal ICA cleaned greyplot from task subject 2 from Power (2019) compared with the task 'on' versus 'off' design convolved with the hemodynamic response function. Power's regions of concern are highlighted with |_ _ _ _| (and also accompanied by a series of vertical tick marks). Task positive regions are located below row ~1000 in the greyplot (see original Figure 13 bottom row blue regions).

representing agreed upon global noise that are completely removed by tICA cleanup, illustrating that the features highlighted by Power are only sometimes temporally coincident with global noise. iii) tICA cleanup is highly comparable to GSR in the greyplots of many epochs that were highlighted in Power's response article. These post-GSR greyplots were included in our original article but omitted from Power's critique. iv) These debated features are removed when sleep-associated components are also removed from the resting-state fMRI data, showing that tICA can successfully separate the features in question from the data by using a different classification.

Second, we disagree that widespread use of manually curated respiratory traces is practical. While we agree that such data can be helpful in aiding classification of components, manually curated traces are not available for the HCP subjects or many other datasets in the community. Moreover, curation of such data requires specialized expertise and would be time consuming, particularly for those with less experience with respiratory trace analyses than Power. For these reasons, curation and quality control of the HCP respiratory data has yet to be completed by the HCP or anyone else, though we expect that the Connectome Coordination Facility (CCF) would be willing to host such data if Power or anyone else was able to generate it. That said, we believe that such an effort is unlikely to change the main conclusions of our paper. Errors in peak detection in respiratory traces are unlikely to increase correlation between physiological noise traces and the fMRI data, and we used those traces that had the highest correlation with the fMRI data – a form of automated quality control. We would of course welcome better automated approaches for review and cleanup of physiological noise traces, as manual curation of such data is impractical for many fMRI studies, particularly large ones. Overall, we believe the best way forward is implementation of an automated classifier for tICA components that is capable of working without explicit need to collect and manually curate respiratory traces.

Third, Power states that we lack a "neural record" to anchor our decisions on component classification. While we do not have an independent record of neural activity, for half of the data presented in the original article we have experimentally manipulated the fMRI signal with a task, a widely accepted method to manipulate neural-activity-related BOLD signals in humans. In Section 2 of this article's Supplement, Supplementary Fig. 4 expands an analysis to all task contrasts that we carried out for just one task contrast in the original article (its Supplementary Fig. 25), showing that compared to global signal regression, tICA cleanup avoids removing spatial patterns of fMRI signal change that mimic the task activation map while (as we showed previously) removing biases from stimulus-correlated respiration, and improving statistical sensitivity to task effects (our original Fig. 6). Interestingly, we find analogous results for movement regressors versus spatial ICA + FIX automated classification for the cleanup of motion artifacts (see Supplementary Figs. 5 and 6 of this article). Thus, we again emphasize the importance of using task fMRI as a positive control for 'benchmarking' BOLD denoising approaches and note that prior studies have generally not incorporated this vital methodological practice, which enriches the signal/artifact classification decisions that we discuss below.

1.2. Classification of components

Although Power did not identify specific components in his critique, we believe based on further direct communication that his main concern relates to our classification of certain components as "neural signal" – specifically, the semi-global sleep-related components RC1 and RC5, and a set of somatotopically organized "DVARS Dips"-associated components (listed below) that have clear sensori-motor spatial activations.

Component RC1 is the strongest resting-state component, but is not present at all in the task data (unlike the respiratory noise components that are present in both task (TC1 and TC30) and resting state (RC3, RC6, and RC8)). Importantly, the spatial structure of RC1 has both semi-global positive and area-specific negative features – something that is not straightforwardly compatible with an effect of global blood flow modulation (see below). It shows strong spatial similarity to patterns seen in previous arousal studies (Horovitz et al., 2008; Tagliazucchi and Laufs, 2014), and its amplitude increases throughout a run as drowsiness/sleepiness is likely increasing for many subjects. Interestingly, its amplitude (per run in each subject) correlates with those of a large group of other neural components (original Supplementary Fig. 27) – all of which may be more active during drowsiness or sleep. This stands in contrast to a second group of neural components that may be more active during the awake state. Importantly, this component is not strongly correlated with RVT (Respiration Volume per Time) or increased in amplitude during DVARS Dips (a more precise measure of head motion, see original Fig. 7 and original Supplementary Figs. 1–3). Based on this information we classified RC1 as a sleep-related neural component. Though investigators interested in awake resting state fMRI may justifiably wish to avoid analyzing epochs of time when this component's amplitude is high (i.e., subjects are likely drowsy or sleeping), regressing it out will likely not adjust for all the other neural changes that accompany drowsiness/sleep (as discussed in the original article). This component's amplitude could be used as an index to identify subjects that were likely to be sleeping so that those subjects could be excluded from analyses that require awake subjects.

Component RC5 is a pan-sensori-motor component that behaves similarly to RC1 in the resting-state, though it is weakly present in task data (TC23) where it is somewhat stronger during DVARS Dips (original Fig. 3). Its functional significance is currently unknown, though its lack of correlation with RVT suggests that it is not a respiratory noise component, and its respect for sensori-motor areal boundaries suggests that it is likely neural in origin.

The most puzzling criticism of our classification relates to the somatotopically organized sensori-motor components RC27/TC16 (Face), RC33/TC6 (Right Upper Extremity), RC39/TC29 (Eye and Trunk), RC40/TC31 (Left Upper Extremity), and RC48/TC19 (Lower Extremities). Hypothetically, if an investigator were to remove these

components from the data, very little somatotopic sensori-motor neural signal would remain, which would, for example, eliminate the neural signal of the HCP's motor task, as the spatial correlations between these components and typical task fMRI GLM beta maps were very high (original Supplementary Figs. 7–10). Indeed these components are highly spatially similar in task vs. resting state, and during the task they show highly specific blocked task driven activity (original Supplementary Figures TC6, TC16, TC19, and TC31). They provide an excellent example of the principle that being temporally correlated with a nuisance indicator (in this case these components are stronger during DVARS Dips, original Fig. 3) does not necessarily make components noise, especially if there is a neurobiological reason for this correspondence. These components are the neural signatures of moving the face, eyes/trunk, left upper extremity, right upper extremity, and lower extremities. Further, we were puzzled by certain remarks in Power's commentary that seemed to imply that motion and respiration are strongly linked, when his previous work has robustly shown them to have differing mechanistic effects on the fMRI signal (Power et al., 2018). While Power also shows case examples supporting his position, we found that motion and respiration are not necessarily closely linked and can be strongly dissociated – e.g., resting state subject 1 (see also below) and subject 2 have high respiratory and global signal but few DVARS Dips (13/1200 frames for subject 1 and 9/1200 frames for subject 2). Thus, we want to highlight that motion, respiratory physiology, and arousal all have different impacts on the fMRI signal through different mechanisms and should be thought of and treated as separate entities, rather than being all lumped together as "motion-related" (even though they can at times have some correlation – e.g., a deep breath may be accompanied by movement of the head; or respiration depth may change as arousal changes).

1.3. Final thoughts

Even if one accepts Power's interpretation of the tICA-driven component classification and greyplots, it is important to ask "what is the magnitude of this effect"? Only one of the example subjects in Power's commentary has physiological noise traces, but by chance that subject is among those with the strongest relationship between RVT and the fMRI data (resting-state subject 1). We find that with no cleanup at all, RVT explains 31.3% of the global timecourse variance in this subject. This drops to 19.7% after sICA + FIX and 3.5% after tICA cleanup. Averaged over the resting state and task runs with RVT traces that had the top 10% correlation with fMRI data, the same values were 13.7%/16.6% without cleanup, 7.8%/9.6% following sICA + FIX, and 1.4%/2.2% following tICA cleanup. As noted above and in this reply (Fig. 1), it is entirely reasonable for respiration to have some correlation with neural activity (indeed we showed a relatively strong correlation of $r = 0.56$ between the blocked task design and RVT in the original paper in Supplementary Fig. 22). As we discussed in the original paper's Supplementary Discussion, the null hypothesis of no correlation between respiration and neural activity is both statistically unlikely and experimentally shown not to occur in the task setting. Thus, we disagree that residual greyplot structure in some cortical areas that is sometimes temporally coincident with removed global respiratory noise indicates incomplete cleanup. To reiterate, the observation that respiratory and neural signals are sometimes temporally coincident (though having differing spatial patterns) does not mean that they arise from the same mechanism – rather we would argue that two distinct processes with some temporal overlap better explain such findings (see this article's supplement and preliminary point 2 above).

Finally, Power argues that he expects the global respiratory signal to have a non-uniform spatial pattern that "follows a sensorimotor distribution," based on prior literature. Though we are aware of the older literature cited by Power (Birn et al., 2006, 2008; Wise et al., 2004), these studies were analyzed using 'legacy' approaches in small samples that (1) relied on volume-based smoothing and volumetric cross-subject alignment that will alter the locations of apparent boundaries in the data (Coalson et al., 2018), and (2) in some cases relied on statistical thresholding that obscures below-threshold information and potentially enhances apparent differences across brain regions (Glasser et al., 2016). This lack of spatial precision leaves considerable fine scale ambiguity as to the hypothesized spatial topography of the global respiratory signal noted by Power. For example, it is not possible to tell whether the patterns appearing in this literature show the abrupt change from positive to negative when entering areas POS2 and RSC seen in the sleep-related component RC1 but not in the respiratory-related components RC3, RC6, RC8, or TC1 – components that at a coarse level will all look relatively similar. Overall, precise, unthresholded spatial maps are critical to resolving the current debate, as data analyzed using legacy techniques may be unable to distinguish these global and semi-global components either visually or quantitatively. Moreover, other legacy studies, after physiological noise cleanup and/or GSR, attributed the visual/auditory/sensori-motor pattern to arousal (noted above and in the original paper; Horovitz et al., 2008 and Tagliazucchi and Laufs, 2014), leaving the question of the precise spatial distribution of the global respiratory signal open, if one does not believe that we fully characterize it with temporal ICA in a group of 449 precisely aligned, unsmoothed, and unthresholded subjects. Regardless of this legacy literature and its methodological barriers to interpretation, we invite Power to propose and test a mechanistic explanation of how changes in overall blood pool pCO2 from respiration would lead to modulation of the fMRI signal in specific cortical areas or subregions that could produce the areal or subregional boundary dependent distributions seen in the above debated components, as opposed to the largely T2* dependent distributions seen in the respiratory tICA components (e.g., TC1, RC6, and RC8), or a vascular distribution seen in other noise components. We are not aware of any convincing theoretical or experimental basis for overall blood pool pCO2 causing *area-specific* BOLD fluctuations, particularly ones that are anti-correlated with the main semi-global fluctuations (e.g., in POS2 and RSC in sleep-related component RC1).

Critically, rigorous experimental testing of this hypotheses would require using analysis approaches that avoid volumetrically blurring the data, accurately align cortical and subcortical areas across subjects, and do not obscure the spatial topography of the entire map with statistical thresholding (Glasser et al., 2016; Coalson et al., 2018). At present, we find that the available evidence best supports a neural origin of these components (RC1, RC5/TC23, RC27/TC16, RC33/TC6, RC39/TC29, RC40/TC31, and RC48/TC19) and thus the conservative approach is to retain them unless and until a neurobiologically sound and experimentally-grounded basis for their removal is presented. Again, we welcome experimental tests that would help the field improve the tICA approach. Such future studies will hopefully further clarify the mechanistic interactions between neural signals, drowsiness/sleep, respiration, and head motion on the fMRI signal. The tICA approach is well positioned as a platform for further investigation of these issues as new experimental evidence is generated, which will be vital for refining temporal ICA component classification. Collectively, we welcome feedback and see room for future studies and improvements, but emphasize that the proposed sICA + tICA framework removes many well-established artifacts (e.g., from movement and respiration) while minimizing the removal of neural signal. In doing so, we believe that this framework provides the most selective, effective, and refinable denoising approach currently available for fMRI data, and thus constitutes a major improvement over GSR and related approaches. Future work will be directed towards automation of the tICA approach to make it widely accessible.

Acknowledgements

Supported in part by the Human Connectome Project, WU-Minn-Ox Consortium (1U54MH091657) funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research; the McDonnell Center for Systems Neuroscience at Washington University;

and NIH F30 MH097312 (M.F.G.), RO1 MH-60974 (D.C.V.E.). Funding to SS, JB, SH gratefully acknowledged via Wellcome Trust strategic award 098369/Z/12/Z. SH was supported by the grant #2017-403 of the Strategic Focus Area "Personalized Health and Related Technologies (PHRT)" of the ETH Domain.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.neuroimage.2019.04.046.

References

Birn, R.M., Diamond, J.B., Smith, M.A., Bandettini, P.A., 2006. Separating respiratory-variation-related fluctuations from neuronal-activity-related fluctuations in fMRI. Neuroimage 31, 1536–1548.

Birn, R.M., Smith, M.A., Jones, T.B., Bandettini, P.A., 2008. The respiration response function: the temporal dynamics of fMRI signal fluctuations related to changes in respiration. Neuroimage 40, 644–654.

Burgess, G.C., Kandala, S., Nolan, D., Laumann, T.O., Power, J.D., Adeyemo, B., Harms, M.P., Petersen, S.E., Barch, D.M., 2016. Evaluation of denoising strategies to address motion-correlated artifacts in resting-state functional magnetic resonance imaging data from the human connectome project. Brain Connect. 6, 669–680.

Coalson, T.S., Van Essen, D.C., Glasser, M.F., 2018. The impact of traditional neuroimaging methods on the spatial localization of cortical areas. Proc. Natl. Acad. Sci. Unit. States Am. 115. E6356-E6365.

Glasser, M.F., Coalson, T.S., Bijsterbosch, J.D., Harrison, S.J., Harms, M.P., Anticevic, A., Van Essen, D.C., Smith, S.M., 2018. Using temporal ICA to selectively remove global noise while preserving global signal in functional MRI data. Neuroimage 181, 692–717.

Glasser, M.F., Smith, S.M., Marcus, D.S., Andersson, J.L., Auerbach, E.J., Behrens, T.E., Coalson, T.S., Harms, M.P., Jenkinson, M., Moeller, S., Robinson, E.C., Sotiropoulos, S.N., Xu, J., Yacoub, E., Ugurbil, K., Van Essen, D.C., 2016. The Human Connectome Project's neuroimaging approach. Nat. Neurosci. 19, 1175–1187.

Horovitz, S.G., Fukunaga, M., de Zwart, J.A., van Gelderen, P., Fulton, S.C., Balkin, T.J., Duyn, J.H., 2008. Low frequency BOLD fluctuations during resting wakefulness and light sleep: a simultaneous EEG-fMRI study. Hum. Brain Mapp. 29, 671–682.

Power, J.P., 2019. Temporal ICA has not properly separated global fMRI signals: a comment on Glasser et al.,., 2018. Neuroimage.

Power, J.D., Mitra, A., Laumann, T.O., Snyder, A.Z., Schlaggar, B.L., Petersen, S.E., 2014. Methods to detect, characterize, and remove motion artifact in resting state fMRI. Neuroimage 84, 320–341.

Power, J.D., Plitt, M., Gotts, S.J., Kundu, P., Voon, V., Bandettini, P.A., Martin, A., 2018. Ridding fMRI data of motion-related influences: removal of signals with distinct spatial and physical bases in multi-echo data. Proc. Natl. Acad. Sci. U. S. A 115, E2105–E2114.

Power, J.D., Plitt, M., Laumann, T.O., Martin, A., 2017. Sources and implications of whole-brain fMRI signals in humans. Neuroimage 146, 609–625.

Tagliazucchi, E., Laufs, H., 2014. Decoding wakefulness levels from typical fMRI resting-state data reveals reliable drifts between wakefulness and sleep. Neuron 82, 695–708.

Wise, R.G., Ide, K., Poulin, M.J., Tracey, I., 2004. Resting fluctuations in arterial carbon dioxide induce significant low frequency variations in BOLD signal. Neuroimage 21, 1652–1664.