Multi-task batteries for individual brain mapping: Experimental design and implementation

Bassel Arafat, Caroline Nettekoven, Jörn Diedrichsen (jdiedric@uwo.ca)
Department of Computer Science and Western Institute of Neuroscience, Western University
London, Ontario, N6A 3K7, Canada

Abstract

While the characterization of individual human brain organization with functional magnetic resonance imaging (fMRI) has in the past relied heavily on resting-state data, it has been shown that a more powerful identification of functional brain organization can be achieved with batteries including a broad set of tasks. Following practical considerations, these multi-task datasets are often designed such that each imaging run includes only a small number of similar tasks or conditions, such that most task-task contrasts have to be made across fMRI runs. Here we show that a design in which all tasks are measured repeatedly within the same imaging run is statistically superior both for estimating tasks-rest contrasts, as well as any task-to-task contrast. An interspersed multi-task design leads to more predictive brain parcellations and connectivity models, even though the design requires participants to constantly switch between tasks. We present a flexible Python toolbox that implements 20+ common tasks with this design, and that automatizes the generation of multi-task batteries for fMRI experimentation. Furthermore, we provide a framework for sharing and integrating pre-processed data across a growing number of multi-task datasets.

Keywords: Multi-task batteries; functional magnetic resonance imaging (fMRI); functional localization; precision mapping; data sharing

Introduction

Functional precision mapping aims to characterize the individual spatial layout and co-activation patterns of different functional regions of the human brain across a large number of task states. While such characterization can be achieved with resting-state data (Tavor et al., 2016), individual brain organization measured with a broad battery of tasks scanned for the same amount of time better generalizes to new tasks (Nettekoven, CCN, 2025). Such multi-task batteries have been shown to be useful in enabling an interpretable and stable mapping of different brain structures (King, Hernandez-Castillo, Poldrack, Ivry, & Diedrichsen, 2019).

Datasets that study individual subjects across many different tasks (deep-phenotyping datasets) usually follow one of two different designs. In the blocked design, adopted for example in the task-fMRI dataset of the Human Connectome project (Barch et al., 2013) or the Individual Brain Charting project (Pinho et al., 2018), each imaging run includes only a subset of tasks or conditions and the rest baseline, with different imaging sessions dedicated to different types of tasks. In the interspersed design, used for example in the Multi-domain task battery dataset (King et al., 2019), each run contains a short period of each task, with every imaging run repeating the same tasks in a different order.

In this paper we compare these two designs. We carefully characterize the variance and covariance of measurement noise from a range of empirical multi-task studies, and then use optimal design calculations (Dale, 1999) to compare a wide range of different designs. We show the advantages of interspersed design for measuring both task-task and task-baseline contrasts.

We present an open-source Python toolbox that enables the effortless design and implementation of Multi-task batteries for fMRI studies. We also present an analysis and data-sharing framework to enable the integration of many multi-task datasets, which will accelerate the development of more complex models of organization of the human brain.

Methods and Results

Sources of measurement noise

To determine the optimal fMRI design for a Multi-task battery, we need to consider two different sources of measurement noise that impact the estimation of task-related activity: First, there is variability due to the measurement of the task itself (σ^2_ε), which for randomized designs is approximately independent across different tasks / conditions within a run. Secondly, measurement noise also affects the common baseline within each run. This noise source (σ^2_b) induces positive covariance between activity estimates of different tasks / conditions of the same run. Both noise sources can for example be clearly seen in the covariance matrices for the activity estimates of 2 runs of the Human Connectome task dataset (Fig. 1a). The condition-wise variance and covariance within a run is substantially higher than the covariance between runs. The between-run covariance allows us to estimate the signal variance (σ^2_s) and covariance (γ_s), such that we can quantify the importance of the two different noise sources across datasets. The noise of the baseline measurement accounted for between 22.1% (Working Memory) and 86.7% (Language) of the total measurement noise, with the other sessions of the HCP dataset falling between these two values.

Interspersed vs. Blocked design

To understand which experimental design is most efficient, we conducted a theoretical analysis with 2 imaging runs and 4 tasks. In the blocked design (Fig. 1b), two tasks were scanned in run 1, and the other two tasks in run 2. In the interspersed design, both runs included all the tasks. The total length of

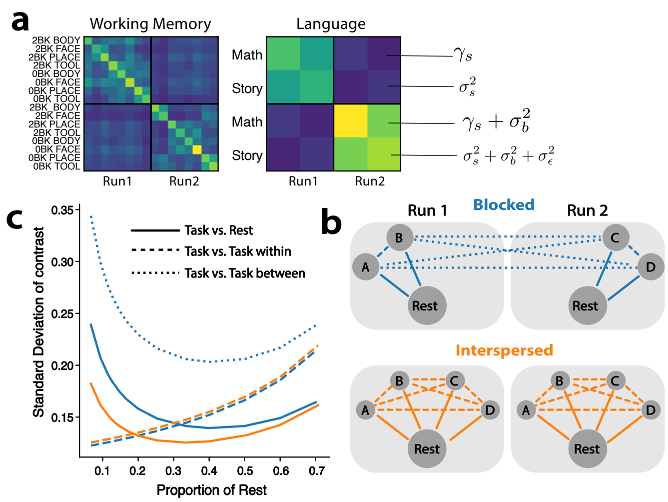

Figure 1: Experimental Design of Multi-task batteries. **(a)** Covariance of cortical activity patterns across two imaging runs for two sessions of the HCP task dataset. The comparison of within-run and between-run covariances allows us to estimate the two noise sources σ_b^2 and σ_ε^2. **(b)** In a blocked design, only some task-task contrast can be made within-run (dashed line), while most contrasts have to be made between runs (dotted line). In a interspersed design, all contrasts can be made within run. **(c)** Predicted standard deviation of task-rest, and task-task contrasts for blocked (blue) and interspersed (orange) designs as a function of the proportion of each imaging run dedicated to rest.

each run was kept constant across the two designs. We also varied the proportion of each run that was spent on measuring the resting baseline. In these designs, the noise on the common baseline accounted for 17.6% (0.7 baseline) and 81.8% (0.1 baseline), showing that the simulations covered a realistic range of scenarios.

Because the same amount of time is spent on each task in the two designs (either within a single run or across two runs), the within-run task-task contrasts (Fig. 1c, dashed line) are measured with the same variability in both designs. As resting baseline is not important here, the lowest variability would be achieved in a design that does not include rest at all. In contrast, the between-run task-task contrasts in the blocked design (dotted line) rely on the common baseline measurement across the two runs. As it is impacted by the measurement noise on the baseline for both run 1 and 2, its variance is more than twice as high as compared to the within-run contrast. Maybe surprisingly, the interspersed design also results in less variable estimates for the (within-run) contrasts between task and rest (solid line). This is due to the fact that in the blocked design, the contrast can only be calculated against the resting baseline of one run, while the interspersed design can leverage the resting baseline in both runs.

Furthermore, we show that the advantage of the interspersed design remains for larger number of tasks, and when we take into account the time lost by switching between tasks within runs. Consistent with earlier papers (Friston, Zarahn, Josephs, Henson, & Dale, 1999), we show that 30s-40s per task within each run provides a good balance between the low-pass properties of the hemodynamic response function and the increased physiological noise at very low frequencies. With this timing and a run length of 12 min, we can measure task batteries with up to 20 different tasks. Such design can be shown to lead to a powerful characterization of the task-evoked activities in individual brains (King et al., 2019).

A Python toolbox for running Multi-task batteries

We are presenting an open-source python toolbox for the design of Multi-task batteries. Built upon PsychoPy (Peirce et al., 2019), the toolbox implements stimulus presentation, response collection, recording of behavioral and eye-tracking data, scanner synchronization, and instructions. With 20+ tasks currently implemented, the toolbox allows for the flexible and fast assembly of new batteries. With an object-oriented design, the toolbox can be easily extended with new tasks and response devices.

Functional Fusion: A call for open source initiatives for integrating different multi-task datasets

The multi-task battery approach is especially powerful when multiple datasets spanning various functional domains can be integrated. To ease such analysis, we established a data management framework that utilizes current BIDS-derivative standards. Currently, we have 10 multi-task datasets, including the Multi-domain task battery (King et al., 2019), the individual brain charting project (Pinho et al., 2018), the Human Connectome task dataset (Barch et al., 2013), and the 103 task dataset (Nakai & Nishimoto, 2020), processed in the framework. The python package enables the quick extraction of functional contrasts and time series in any desired group atlas space . It also provides statistical tools for reliability analysis, and the alignment of different task batteries into a common library. The latter process is especially powerful when different task-batteries contain a number of shared "anchor" tasks, a feature enabled by the standardized task framework.

Summary

The study of individual human brains is greatly aided by the application of broad battery of tasks, which provide a robust, generalizable, and interpretable characterization of functional brain organization. To accelerate the development of more sophisticated models of individual brain organization, we are inviting other research groups to contribute their own multi-task dataset to the growing collection. Each participating laboratory will have free access to pre-processed datasets that can be analyzed using a unified analysis framework, while retaining control of the rules for sharing their own datasets.

Acknowledgments

This work was supported by a Discovery Grant from the Natural Sciences and Engineering Research Council of Canada (NSERC, RGPIN-2016-04890), and a project grant from the Canadian Institutes of Health Research (CIHR, PJT-507612), both to J.D. Additional funding came from the Canada First Research Excellence Fund (BrainsCAN) to Western University.

References

Barch, D. M., Burgess, G. C., Harms, M. P., Petersen, S. E., Schlaggar, B. L., Corbetta, M., ... Van Essen, D. C. (2013). Function in the human connectome: task-fMRI and individual differences in behavior. *Neuroimage*, *80*, 169–189.

Dale, A. M. (1999). Optimal experimental design for event-related fMRI. *Hum. Brain Mapp.*, *8*, 109–114.

Friston, K. J., Zarahn, E., Josephs, O., Henson, R. N., & Dale, A. M. (1999). Stochastic designs in event-related fMRI. *Neuroimage*, *10*(5), 607–619.

King, M., Hernandez-Castillo, C. R., Poldrack, R. A., Ivry, R. B., & Diedrichsen, J. (2019, August). Functional boundaries in the human cerebellum revealed by a multi-domain task battery. *Nat. Neurosci.*, *22*(8), 1371–1378.

Nakai, T., & Nishimoto, S. (2020, December). Quantitative models reveal the organization of diverse cognitive functions in the brain. *Nat. Commun.*, *11*(1).

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., ... Lindeløv, J. K. (2019, February). PsychoPy2: Experiments in behavior made easy. *Behav. Res. Methods*, *51*(1), 195–203.

Pinho, A. L., Amadon, A., Ruest, T., Fabre, M., Dohmatob, E., Denghien, I., ... Thirion, B. (2018, June). Individual brain charting, a high-resolution fMRI dataset for cognitive mapping. *Scientific data*, *5*.

Tavor, I., Parker Jones, O., Mars, R. B., Smith, S. M., Behrens, T. E., & Jbabdi, S. (2016). Task-free MRI predicts individual differences in brain activity during task performance. *Science*, *352*(6282), 216–220.