



OPEN

DATA DESCRIPTOR

# An fMRI dataset for investigating language control and cognitive control in bilinguals

Tingting Guo<sup>1</sup>, Xuedi Liu<sup>1</sup>, Mo Chen<sup>2</sup>, Yongben Fu<sup>3</sup> & Taomei Guo<sup>1</sup>✉

The neural correlation between language control and cognitive control in bilinguals remains an area ripe for further exploration. In this work, we present a functional magnetic resonance imaging (fMRI) dataset that simultaneously examines both types of control in 77 healthy, unrelated Chinese-English bilinguals. Each participant completed a language switching task to assess language control and a rule switching task to evaluate cognitive control while undergoing functional MRI scanning. We collected structural imaging data, task-related functional imaging data, and behavioral data from the participants. Additionally, their language proficiency and domain-general cognitive ability were recorded after the scanning. This dataset was released early to facilitate the exploration of the neural relationship between language control and cognitive control, promoting its broader use and benefiting the scientific community. It is well-suited for the study of causality analysis, representational similarity analysis and the prediction model construction.

## Background & Summary

For bilinguals, cross-language competition during language production calls for the involvement of language control<sup>1</sup>. It has been shown that there is a close relationship between language control and domain-general cognitive control<sup>2–4</sup>. Over the past decade, many neuroimaging studies have attempted to examine the neural overlapping and dissociation between these two types of cognitive processes<sup>5–7</sup>. Some of our studies have revealed how cortical and subcortical brain regions coordinate to work for language control and cognitive control<sup>7</sup>.

However, with the ongoing deepening and diversification of data analysis methods, there is still great potential for further exploration of the relationship between language control and cognitive control. To support this endeavor, we present an fMRI dataset capturing the neural processing of 77 Chinese-English bilinguals as they perform language control and cognitive control tasks. In the language control task, participants were required to name pictures in either their native language (Chinese) or their second language (English) based on given cues. In the cognitive control task, participants needed to respond to the direction of arrows according to different rules indicated by cues. The two tasks were carefully balanced in terms of sequence, procedure design, and experimental environment to facilitate direct comparison of the two types of brain processing. In addition to structural and functional MRI data, behavioral data and relevant ability scores were also collected.

The current dataset includes data from 77 participants in our previous studies<sup>7–9</sup> with each addressing distinct research questions using different analytical approaches: Chen *et al.*, 2019 utilized data from these participants' cognitive control tasks, focusing on group differences (experimental vs. control) by comparing brain activation between incongruent and congruent conditions using General Linear Model (GLM) methods; Wu *et al.*, 2019 investigated differences in effective connectivity networks between two tasks by using the extended unified structural equation modeling (euSEM) approach to extract time series based on event onsets; Yuan *et al.*, 2021 used data from both tasks of these participants, applying euSEM to extract time series and examine the effects of task order on effective connectivity.

Although we have previously published results from a subset of participants in this dataset<sup>7–9</sup>, the raw and unprocessed nature of the shared dataset offers ample opportunities for reuse and extension, enabling researchers to apply diverse fMRI data analysis methods tailored to their specific research questions. For example,

<sup>1</sup>State Key Laboratory of Cognitive Neuroscience and Learning & IDG/McGovern Institute for Brain Research, Beijing Normal University, Beijing, China. <sup>2</sup>School of Psychological and Cognitive Sciences, Peking University, Beijing, China. <sup>3</sup>The Psychological Education and Counseling Center, Huazhong Agricultural University, Wuhan, China. ✉e-mail: [quotm@bnu.edu.cn](mailto:quotm@bnu.edu.cn)

future work using this dataset could employ representational similarity analysis or multivoxel pattern analysis to assess similarities between tasks<sup>10–12</sup>. Causal analyses, such as granger causality analysis or dynamic causal modeling<sup>13–15</sup>, are also well-suited to these data, allowing for exploration of the dynamic functional connectivity between language and cognitive control. Additionally, by integrating machine learning algorithms, connectome-based prediction models could be developed to predict psychological states from physiological data<sup>16,17</sup>. The structural images can also be combined with functional imaging or/and behavioral data to offer novel neurobiological insights into the relationship between language and cognitive control<sup>18,19</sup>. Moreover, the dataset is formatted in the brain imaging data structure (BIDS)<sup>20</sup> and is publicly available on OpenNeuro.org<sup>21–23</sup>, making it easily integrable with other datasets for large-scale meta-analyses or cross-study comparisons. In summary, the dataset's rich and diverse data, combined with its compatibility with modern analytical techniques, provides a valuable resource for advancing research in this field. We hope this publication will inspire further exploration and benefit the scientific community.

## Methods

**Participants.** A total of 77 Chinese-English bilinguals (34 females), aged 19 to 30 years ( $M = 22.2$  years,  $SD = 2.2$  years) with normal general cognitive ability (Raven score range = 44–60,  $M = 56.19$ ,  $SD = 3.66$ ) participated in the present study. They were all right-handed with normal or corrected-to-normal vision, and reported no psychiatric or neurological disorders. All participants were recruited by posting fliers on campus and were paid for their participation (CNY¥100 per hour).

**Language ability assessments.** They were all native Chinese speakers and began to learn English as a second language around the age of 9.4 years ( $SD = 2.6$  years). They all passed the College English Test Band 4 (CET-4;  $M = 521.45$ ,  $SD = 51.85$ , full score = 710), an obligatory normalized English test for college students in China. According to their self-rating scores for both languages on a 10-point scale for listening, speaking, reading, and writing, their proficiency in Chinese ( $M = 8.40$ ,  $SD = 1.15$ ) was higher than in English ( $M = 5.65$ ,  $SD = 1.25$ ;  $t(76) = 16.27$ ,  $p < 0.001$ ), indicating that they were relatively proficient, but Chinese-dominant bilinguals.

It should be noted that the CET-4 score is based on a norm-referenced approach. The norming group consists of approximately 30,000 non-English major candidates selected from 16 universities across China. The raw scores from each examination, after equating, are converted into reported scores using a norming formula. According to information disclosed by the CET-4 organizing authority, the National Education Examinations Authority of the Ministry of Education of China (<https://cet.neea.edu.cn/html1/folder/19081/5124-1.htm>), the normative percentile ranks corresponding to different score ranges are shown in Table S1 (see Supplementary Information document).

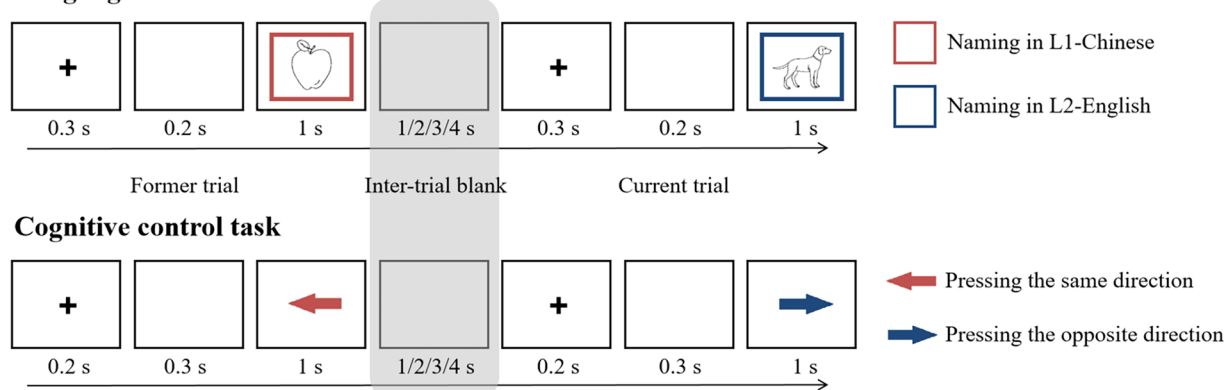
**Ethical approval.** All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards. All participants gave their written informed consent to participate in the study and to publicly share their de-identified data. Before the experiment, they were given sufficient time to carefully read and sign the informed consent form. They were also encouraged to consult our research team with any questions or concerns, ensuring full understanding of the study's terms and procedures. This study was approved by the Institutional Review Board of the Imaging Center for Brain Research of Beijing Normal University under Protocol Number ICBIR\_A\_0012\_003.

**Procedures.** All participants performed the language control task and the cognitive control task in the MRI scanner after a brief practice. The order of tasks was counterbalanced across participants. The formal experiment in the scanner included two runs for each task. Each run lasted 5.5 minutes. An 8-minute anatomical scan was applied afterward. Participants were allowed to take a break between runs, thus the whole scanning lasted about 40 minutes. Each run consisted of 82 trials, with the first two trials added as fillers to eliminate data from the statistical analyses for the instability of the magnetic field. Thus each of the two tasks contained 160 formal trials.

Due to technical limitation, we did not record participants' naming responses in the scanner. To obtain their naming data for language control performance metrics, participants repeated the language control task in a behavioral laboratory after MRI scanning. The laboratory session for the language control task was conducted in a sound-attenuated room using the same task paradigm as in the scanner, ensuring consistency in task demands while allowing for precise behavioral data collection. After that, participants also completed a questionnaire on language learning history and proficiency and then performed the Raven's Standard Progressive Matrices task<sup>24</sup> to measure their general cognitive ability<sup>25,26</sup>.

**Tasks.** The language switching task and rule switching task were used to examine the neural correlates associated with language control and cognitive control, respectively. Thus, we would consistently refer to the two tasks as the language control task and the cognitive control task in the following. Stimuli were presented via E-Prime 2.0 (RRID:SCR\_009567, Psychology Software Tools, Pittsburgh, PA).

*The language control task.* Forty-eight line drawings of common objects were selected from the database of Snodgrass and Vanderwart<sup>27</sup> as stimuli for the task. Eight of these drawings were used as practice trials or fillers. Participants were instructed to name a picture in either Chinese or English according to the cue (i.e., a red or blue frame of the picture). Specifically, if the frame was red, participants needed to name the picture in Chinese; if the frame was blue, they needed to name the picture in English (see Fig. 1). They were asked to name the picture in a soft voice to eliminate possible head motions. The color-language assignment was counterbalanced across participants. The change in the frame color between the current trial and the former trial indicated that participants needed to switch the language, whereas the same color as the former trial signaled that they should

**Language control task**

**Fig. 1** Schematic illustration of the two tasks. The color-language/rule assignment was counterbalanced across participants. For each participant, the correspondence between the language and the rule was fixed; for example, if red was associated with “naming in Chinese” in the language control task, it also indicated that red corresponded to “pressing the same direction” in the cognitive control task.

maintain the same language. This created switch and non-switch conditions, each containing 80 trials. For each trial, a fixation point was presented for 0.3 s, followed by a blank screen for 0.2 s. Then, a picture was presented at the center of the screen for 1 s. Participants were instructed to respond as quickly and accurately as possible. The inter-stimuli intervals were pseudo-randomly varied to present a blank screen for 1, 2, 3, or 4 s to better estimate the blood-oxygen-level-dependent (BOLD) response in MRI scans.

**The cognitive control task.** Blue and red arrows pointing left or right were used as stimuli for the task. Participants were instructed to respond to each arrow following different rules according to the cue which was the arrow's color (red or blue). Specifically, if the arrow was red, participants needed to press the key on the same side as the arrow's direction; if the arrow was blue, they needed to press the key in the opposite direction (see Fig. 1). The color-rule assignment was counterbalanced across participants. A change in the arrow's color between the current trial and the former trial indicated that participants needed to switch the rule, whereas the same color as the former trial signaled that they should maintain the same rule. This created switch and non-switch conditions, each containing 80 trials. For each trial, a fixation point was presented for 0.2 s, followed by a blank screen for 0.3 s. Then, an arrow was presented at the center of the screen for 1 s. Participants were instructed to respond as quickly and accurately as possible. The inter-stimuli intervals were pseudo-randomly varied to present a blank screen for 1, 2, 3, or 4 s to better estimate the BOLD response in MRI scans.

**Neuroimaging data collection.** Whole brain image data were acquired by a 3-Tesla Trio MRI scanner (Siemens Healthineers, Erlangen, Germany) at the MRI Center of Beijing Normal University. During scanning, participants laid down with their heads fixed to minimize the head motion. In each task session, 164 functional T2\*-weighted echo-planar images (EPI) were acquired using an event-related design with the following parameters: slice number = 33, TR (repeated time) = 2000 ms, TE (echo time) = 20 ms, flip angle = 90°, field-of-view (FOV) = 200 × 200 mm<sup>2</sup>, matrix size = 64 × 64, resolution within slices = 3.125 × 3.125 mm<sup>2</sup>, slice thickness/gap = 4.0/0.8 mm. 144 high-resolution T1-weighted anatomical images were also obtained using the following scan parameters: TR = 2530 ms, TE = 3.39 ms, flip angle = 7°, FOV = 256 × 256 mm<sup>2</sup>, matrix size = 256 × 256, resolution within slices = 1.0 × 1.0 mm<sup>2</sup>, slice thickness = 1.33 mm. All images were interleaved from bottom to top in the whole-brain acquisition.

**Neuroimaging data processing.** DICOMs were first converted to NIfTI form using dcm2nii<sup>28</sup> and organized into BIDS 1.8.0<sup>20</sup> (RRID:SCR\_016124) using a heuristic-centric DICOM converter (HeuDiConv; RRID:SCR\_017427) 1.1.6. To ensure anonymization, all structural data have been defaced using pydeface 2.0.2 (<https://github.com/poldracklab/pydeface>) and the data collection dates have been replaced with a random time.

The preprocessing steps were implemented using the standard pipeline of fMRIPrep 24.0.0<sup>29</sup> (RRID:SCR\_016216), which is based on Nipype 1.8.6<sup>30</sup> (RRID:SCR\_002502). For structural data, the T1w image was corrected for intensity non-uniformity (INU) using the N4BiasFieldCorrection<sup>31</sup> (from ANTs 2.5.1<sup>32</sup>, RRID:SCR\_004757) and served as the T1w reference for subsequent steps. Skull-stripping was performed using the antsBrainExtraction.sh workflow (from ANTs). Brain tissue into cerebrospinal fluid (CSF), white matter (WM), and gray matter (GM) was conducted using FAST<sup>33</sup> (from FSL, RRID:SCR\_002823). Spatial normalization to the standard MNI152NLin2009cAsym space was achieved through nonlinear registration using ants-Registration (from ANTs), applied to both the brain-extracted T1w reference and the T1w template. The ICBM 152 Nonlinear Asymmetrical template version 2009c<sup>34</sup> (RRID:SCR\_008796) was selected for normalization and accessed via TemplateFlow 24.2.0<sup>35</sup>.

For functional data of each participant, preprocessing of four BOLD runs also followed a standardized pipeline. A reference volume was created using a custom fMRIPrep method to facilitate head motion correction. Head-motion parameters, including transformation matrices and six rotation/translation components,

were estimated using *mcflirt*<sup>36</sup> (from FSL) prior to spatiotemporal filtering. The BOLD reference was then co-registered to the T1w reference using *mri\_coreg* (from FreeSurfer) followed by a boundary-based registration<sup>37</sup> cost function through *flirt*<sup>38</sup> (from FSL). Co-registration was performed with six degrees of freedom to optimize spatial alignment.

Following preprocessing, a comprehensive set of potential confounds (or nuisance regressors) was generated to support researchers in selecting the most appropriate denoising strategy for downstream analyses. These included framewise displacement (FD), DVARS, and three region-wise global signals derived from the CSF, WM, and whole-brain masks<sup>39</sup>. Additionally, component-based noise correction (CompCor) was applied, with principal components calculated separately for temporal (tCompCor) and anatomical (aCompCor) variants<sup>40</sup>. For detailed descriptions of the confounds' principles and computational methods, please refer to the fMRIPrep documentation (<https://fmriprep.org/en/stable/outputs.html#confounds>).

**Quality assurance.** For the cognitive control task, behavioral data quality was assessed by examining error trials. A high proportion of error trials would indicate either that the participant failed to successfully complete the task or that unexpected technical malfunctions occurred during data collection. For the language control task, due to technical limitations, verbal responses could not be recorded during scanning. We ensured data validity through experimenter monitoring of real-time task engagement, and behavioral performance verification after scanning.

For the neuroimaging data, a set of image quality metrics (IQMs) were generated using MRIQC 21.0.0rc2<sup>41</sup> (RRID: SCR\_022942). The provided IQMs can be used to exclude outlier runs and subjects or as covariates for higher-level analysis. This is especially useful when there are differences between groups of interest in motion and other aspects of image quality that may induce spurious results. From the IQMs generated by MRIQC, a subset of representative metrics was selected for statistical analysis to ensure data quality and interpretability. For structural images, we evaluated the contrast-to-noise ratio (CNR) and the full-width at half-maximum (FWHM), which are critical indicators of spatial resolution and signal clarity<sup>42–44</sup>. For functional images, we assessed the frame displacement (FD) and the temporal signal-to-noise ratio (tSNR), as these metrics are widely recognized for their relevance to motion artifacts and temporal stability in fMRI data<sup>42–44</sup>.

**Generation of single-trial BOLD response.** For the convenience of potential users, we also provided single-trial estimates of BOLD responses through the Least-Squares Separate (LSS) technique. This technique is a robust signal estimation method commonly used in rapid event-related designs, helping to reduce collinearity between estimates<sup>45–47</sup>.

The BOLD response for each preprocessed image was estimated by the General Linear Model (GLM) with local autocorrelation correction<sup>48</sup> using FMRI Expert Analysis Tool (FEAT) 6.0.0<sup>49</sup>, part of FMRIB Software Library (FSL) 6.0.3. We fitted a separate GLM for each stimulus, with the trial of interest modeled as one regressor and all non-target trials collapsed into another regressor. Each model included data only from the same run. The duration of each regressor was defined as 1 s, matching the stimulus presentation period from onset to offset. All regressors were convolved with a double-gamma hemodynamic response function and high-pass filtered (with the cut-off at 100 s). Confounds from the fMRIPrep outputs was selected for each model, following the single-trial assessment methods described by Smith<sup>43</sup>. These confounds comprised FD, three translation and rotation motion parameters, non-steady state volumes, cosine basis functions for temporal filtering, and the first six anatomical CompCor regressors. No smoothing was applied, and non-brain voxels were removed using the Brain Extraction Tool (BET)<sup>50</sup>. Grand-mean intensity normalization was applied to the 4D NIfTI image by a single multiplicative factor.

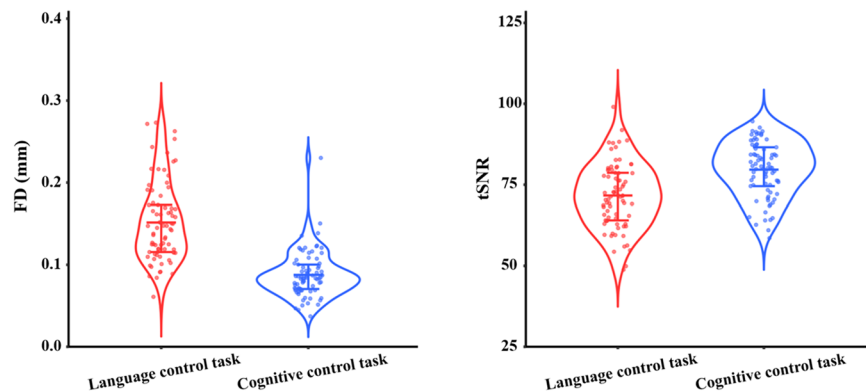
## Data Records

The dataset is publicly available at OpenNeuro ds005455<sup>51</sup> (<https://doi.org/10.18112/openneuro.ds005455.v1.1.5>). All data are structured in compliance with the BIDS standard 1.8.0<sup>20</sup>, which is an increasingly popular framework for describing imaging data in a standardized format. The file named *participants.tsv* contains demographic information and relevant test scores for all participants, as well as the behavioral results of the language control task which were collected in the behavioral laboratory and organized by conditions. Refer to Table S2 (see Supplementary Information document) for the complete data dictionary.

**Raw data.** Each participant's raw data is organized in participant-specific directories with the naming scheme <sub-XXX>. Within each directory, the 'anat' subdirectory contains structural MRI data that has been anonymized through defacing, while the 'func' subdirectory includes the functional data in its raw, unprocessed state. The data are organized using key-value pairs: "sub- <value>" for participant identifiers, "task- <value>" for task types (LanguageControl or CognitiveControl), and "run- <value>" for separate acquisitions under identical parameters, subject, and task conditions. The json files contain information about the acquisition parameters. Event-related data are stored in \*\_events.tsv files, with detailed descriptions provided in the respective data dictionaries: cognitive control task and language control task (see Table S2).

**Derivatives.** The 'derivatives' directory contains three sub-directories: "mriqc", "tsnr" and "singletrial".

**'mriqc' directory.** This directory contains the output from MRIQC, including group-level quality metrics for raw structural and functional images, stored in *group\_T1w.tsv* and *group\_bold.tsv*, respectively. The data dictionaries for these files are summarized in Table S2. Additionally, the directory provides summary reports for structural and functional images in the form of *group\_T1w.html* and *group\_bold.html*, which offer visualizations and statistical summaries of image quality metrics. Participant-specific subdirectories organize files describing raw image quality for each participant, task, and run.



**Fig. 2** Functional data quality metrics by task. Left panel displays framewise displacement (FD) distributions, while right panel shows temporal signal-to-noise ratio (tSNR) distributions. Violin plots represent group-level data distributions, with individual points indicating subject means (averaged across 2 runs per subject). Central horizontal lines mark group medians, and error bars denote interquartile ranges (25th–75th percentiles). Data were extracted from the `group_bold.tsv` file in the MRIQC derivatives directory.

**'tsnr' directory.** This directory includes volumetric images for each participant, task, and run, representing the temporal signal-to-noise ratio (tSNR) computed as the mean signal divided by the standard deviation of the corresponding functional image. Additionally, it contains whole-brain mean tSNR images computed for each task.

**'singletrial' directory.** This directory stores single-trial estimates for each participant, with individual data organized in participant-specific subdirectories. Inside each subdirectory, a 4D NIfTI image represents a single run of a specific task, where each volume corresponds to a trial arranged in chronological order.

### Technical Validation

We performed several procedures to validate our behavioral data, structural and functional neuroimaging data. For each type of data, we presented the results for all participants. For the functional imaging results, we reported comparisons between the two tasks. Although differences between tasks might not be the focus of most studies, other studies may need to control for these differences. We note that no data have been excluded from the published dataset. Thus, users may confirm whether to exclude any data. The results of these validations are as follows.

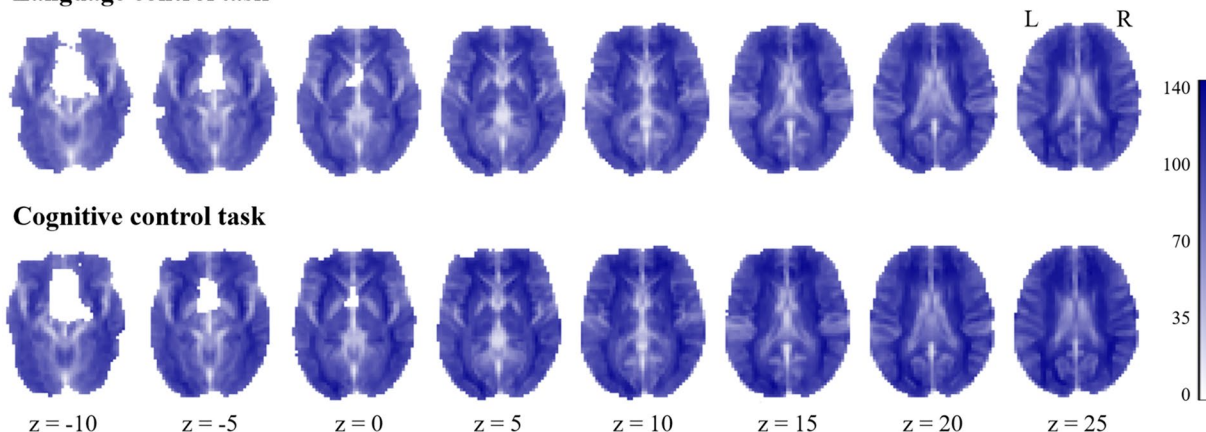
**Behavioural data.** Error rate for each task was examined to ensure sufficient trials for further analyses. For the language control task, behavioral data were collected during a separate laboratory session after scanning. For the cognitive control task, behavioral data were recorded directly during scanning. The switching costs (switch vs. non-switch condition) in response time data were calculated separately for the language control task and cognitive control task.

In both tasks, a relatively low error rate ( $M = 3.99\%$ ,  $SD = 3.42\%$  for the language control task;  $M = 2.72\%$ ,  $SD = 2.50\%$  for the cognitive control task) was observed, indicating that participants were able to complete the tasks successfully. Paired  $t$ -test results for response times revealed that switch trials ( $M = 904$  ms,  $SD = 125$  ms) were significantly longer than non-switch trials ( $M = 864$  ms,  $SD = 114$  ms),  $t(76) = 10.00$ ,  $p < 0.001$  in the language control task; and in the cognitive control task, switch trials ( $M = 606$  ms,  $SD = 106$  ms) were significantly longer than non-switch trials ( $M = 571$  ms,  $SD = 97$  ms),  $t(76) = 10.64$ ,  $p < 0.001$ . These results demonstrated a significant switching cost in both tasks.

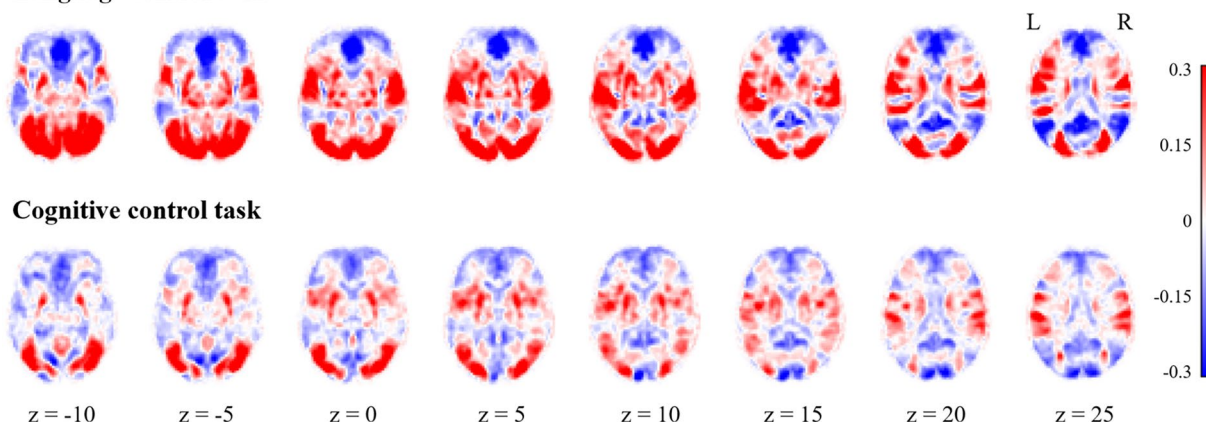
**Structural neuroimaging data.** The quality of neuroimaging images was validated and assessed using MRIQC. We first checked the contrast-to-noise ratio (CNR), an extension of the signal-to-noise ratio (SNR) calculation to evaluate how separated the tissue distributions of grey and white matter are<sup>52</sup>. All participants had high CNR ( $M = 3.89$ ,  $SD = 0.28$ ), indicating clear separation between grey and white matter tissues. Then we examined the full-width at half-maximum (FWHM) of the spatial distribution of the image intensity values in units of voxels<sup>53</sup>, which can serve as a reference for data smoothness. We observed low values for each participant ( $M = 3.63$ ,  $SD = 0.16$ ), indicating that data smoothness was well-maintained. This information can also guide other processing choices, such as the amount of smoothing to apply or whether to smooth to a predetermined level of smoothing<sup>54</sup>.

**Functional neuroimaging data.** During MRI scanning, assessing and minimizing motion is undoubtedly crucial because high amounts of head motion result in low data quality and validity. Therefore, we first checked the extent of head motion in each task, with particular attention to frame displacement (FD), which summarizes the volume-to-volume changes in head position. The FD distribution across tasks is shown in Fig. 2. In both tasks, participants showed relatively low average FD ( $M = 0.15$ ,  $SD = 0.05$  for the language control task;  $M = 0.09$ ,  $SD = 0.03$  for the cognitive control task), indicating very minimal head motions. However, the difference between the two tasks was significant,  $t(76) = 12.00$ ,  $p < 0.001$ , showing that participants exhibited more head motion during the language control task compared to the cognitive control task. Considering that participants needed to



**Language control task**

**Fig. 3** Whole-brain temporal signal-to-noise ratio (tSNR) averaged across all participants for each task. The image displays calculations across all voxels, demonstrating consistently high and spatially uniform throughout both tasks.

**Language control task**

**Fig. 4** Mean single-trial estimates across all participants for each task. Single-trial estimates were generated using the Least Squares - Separate (LSS) technique.

verbally name pictures in the language task, greater head movement is reasonable compared to the cognitive task, which only required button presses, despite our instructions to minimize movement.

Next, we examined the temporal SNR (tSNR) in each task. The tSNR, a commonly used metric for characterizing acquisition performance<sup>55</sup>, can effectively identify data severely compromised by head motion, RF coil issues, or other imaging artifacts; low tSNR indicates low-quality and potentially problematic functional data. The tSNR distribution across task is visualized in Fig. 2. We found that the average tSNR of each participant was high ( $M = 71.66$ ,  $SD = 10.07$  for the language control task;  $M = 79.64$ ,  $SD = 8.62$  for the cognitive control task) and relatively uniform across the whole brain in both tasks (see Fig. 3). A significant difference between two tasks was also found,  $t(76) = -7.32$ ,  $p < 0.001$ , showing that the image quality in the language task was poorer than in the cognitive task. We further conducted correlation analyses between the average FD and tSNR of each participant in two tasks, both of which revealed significant negative correlations:  $r = -0.74$ ,  $p < 0.001$  for language control task; and  $r = -0.41$ ,  $p < 0.001$  for cognitive control task. This indicates that the lower tSNR observed in the language task may be partially due to increased head motion. Note that the tSNR can vary with the imaging parameters (number of averages, resolution, echo time, parallel imaging acceleration, field strength, etc.), making it difficult to set a strict threshold. However, the mean tSNR over the whole brain can be compared with that of other individuals within the group acquired with similar imaging parameters at same site<sup>54</sup>.

Finally, we assessed average task-related activation in each task based on the single-trial estimates. We observed similar activation patterns across the two tasks (Fig. 4). We found that the average response across single trials was associated with increased activation in the visual and motor cortices, which was expected as both tasks required participants to focus on visual stimuli and respond accordingly. We also found that the posterior cingulate cortex and precuneus were deactivated in both tasks, which aligns with previous findings that the default mode network tends to be suppressed when individuals perform cognitively demanding tasks<sup>56,57</sup>.

Additionally, we observed activation in brain regions associated with language production, such as the inferior frontal gyrus and superior temporal gyrus<sup>58,59</sup>, during the language task, indicating that participants engaged in language production, i.e., naming pictures, during the scanning.

### Code availability

All code is openly available on GitHub ([https://github.com/GttNeuro/Guo-Lab\\_datapaper](https://github.com/GttNeuro/Guo-Lab_datapaper)). The scripts related to whole-brain tSNR calculations and single-trial estimations were adapted from code available in the GitHub repository associated with the srndna-datapaper study<sup>43</sup> (<https://github.com/DVS-Lab/srndna-datapaper>). This repository also includes template files used in FEAT and stimulus images for the two tasks described in the main text.

Received: 16 September 2024; Accepted: 19 May 2025;

Published online: 28 May 2025

### References

- Green, D. W. Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition* **1**, 67–81 (1998).
- Bialystok, E., Craik, F. I. M., Green, D. W. & Gollan, T. H. Bilingual Minds. *Psychol Sci Public Interest* **10**, 89–129 (2009).
- Morales, J., Gómez-Ariza, C. J. & Bajo, M. T. Dual mechanisms of cognitive control in bilinguals and monolinguals. *Journal of Cognitive Psychology* (2013).
- Lai, G. & O'Brien, B. A. Examining Language Switching and Cognitive Control Through the Adaptive Control Hypothesis. *Front. Psychol.* **11**, (2020).
- Green, D. W. & Abutalebi, J. Language control in bilinguals: The adaptive control hypothesis. *Journal of Cognitive Psychology* (2013).
- De Baene, W., Duyck, W., Brass, M. & Carreiras, M. Brain Circuit for Cognitive Control Is Shared by Task and Language Switching. *Journal of Cognitive Neuroscience* **27**, 1752–1765 (2015).
- Wu, J. *et al.* Brain network reconfiguration for language and domain-general cognitive control in bilinguals. *NeuroImage* **199**, 454–465 (2019).
- Chen, M. *et al.* Individual differences in inhibitory control abilities modulate the functional neuroplasticity of inhibitory control. *Brain Struct Funct* **224**, 2357–2371 (2019).
- Yuan, Q. *et al.* Neural interaction between language control and cognitive control: Evidence from cross-task adaptation. *Behavioural Brain Research* **401**, 113086 (2021).
- Kriegeskorte, N., Mur, M. & Bandettini, P. A. Representational similarity analysis - connecting the branches of systems neuroscience. *Front. Syst. Neurosci.* **2**, (2008).
- Kragel, P. A. *et al.* Generalizable representations of pain, cognitive control, and negative emotion in medial frontal cortex. *Nat Neurosci* **21**, 283–289 (2018).
- Yuan, Q. *et al.* Patterns and networks of language control in bilingual language production. *Brain Struct Funct* **226**, 963–977 (2021).
- Seth, A. K., Barrett, A. B. & Barnett, L. Granger Causality Analysis in Neuroscience and Neuroimaging. *J. Neurosci.* **35**, 3293–3297 (2015).
- Cai, W. *et al.* Causal Interactions Within a Frontal-Cingulate-Parietal Network During Cognitive Control: Convergent Evidence from a Multisite–Multitask Investigation. *Cerebral Cortex* **26**, 2140–2153 (2016).
- Nee, D. E. & D'Esposito, M. Causal evidence for lateral prefrontal cortex dynamics supporting cognitive control. *eLife* **6**, e28040 (2017).
- Fountain-Zaragoza, S., Samimy, S., Rosenberg, M. D. & Prakash, R. S. Connectome-based models predict attentional control in aging adults. *NeuroImage* **186**, 1–13 (2019).
- Rosenberg, M. D., Hsu, W.-T., Scheinost, D., Todd Constable, R. & Chun, M. M. Connectome-based Models Predict Separable Components of Attention in Novel Individuals. *Journal of Cognitive Neuroscience* **30**, 160–173 (2018).
- Rieck, J. R., Baracchini, G. & Grady, C. L. Contributions of Brain Function and Structure to Three Different Domains of Cognitive Control in Normal Aging. *Journal of Cognitive Neuroscience* **33**, 1811–1832 (2021).
- Berlot, R., Metzler-Baddeley, C., Ikram, M. A., Jones, D. K. & O'Sullivan, M. J. Global Efficiency of Structural Networks Mediates Cognitive Control in Mild Cognitive Impairment. *Front. Aging Neurosci.* **8**, (2016).
- Gorgolewski, K. J. *et al.* The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments. *Sci Data* **3**, 160044 (2016).
- Poldrack, R. A. *et al.* Toward open sharing of task-based fMRI data: the OpenfMRI project. *Front. Neuroinform.* **7**, (2013).
- Poldrack, R. A. & Gorgolewski, K. J. OpenfMRI: Open sharing of task fMRI data. *NeuroImage* **144**, 259–261 (2017).
- Markiewicz, C. J. *et al.* The OpenNeuro resource for sharing of neuroscience data. *eLife* **10**, e171774 (2021).
- Raven, J. C. & Court, J. H. *Raven's Progressive Matrices and Vocabulary Scales*. (Oxford Psychologists Press, 1998).
- Del Missier, F., Mäntylä, T. & de Bruin, W. B. Decision-making Competence, Executive Functioning, and General Cognitive Abilities. *Journal of Behavioral Decision Making* **25**, 331–351 (2012).
- Rohde, T. E. & Thompson, L. A. Predicting academic achievement with cognitive ability. *Intelligence* **35**, 83–92 (2007).
- Snodgrass, J. G. & Vanderwart, M. A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory* **6**, 174–215 (1980).
- Li, X., Morgan, P. S., Ashburner, J., Smith, J. & Rorden, C. The first step for neuroimaging data analysis: DICOM to NIfTI conversion. *Journal of Neuroscience Methods* **264**, 47–56 (2016).
- Esteban, O. *et al.* fMRIPrep: a robust preprocessing pipeline for functional MRI. *Nat Methods* **16**, 111–116 (2019).
- Gorgolewski, K. *et al.* Nipype: A Flexible, Lightweight and Extensible Neuroimaging Data Processing Framework in Python. *Front. Neuroinform.* **5**, (2011).
- Tustison, N. J. *et al.* N4ITK: Improved N3 Bias Correction. *IEEE Transactions on Medical Imaging* **29**, 1310–1320 (2010).
- Avants, B. B., Epstein, C. L., Grossman, M. & Gee, J. C. Symmetric diffeomorphic image registration with cross-correlation: Evaluating automated labeling of elderly and neurodegenerative brain. *Medical Image Analysis* **12**, 26–41 (2008).
- Zhang, Y., Brady, M. & Smith, S. Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. *IEEE Transactions on Medical Imaging* **20**, 45–57 (2001).
- Fonov, V., Evans, A., McKinstry, R., Alml, C. & Collins, D. Unbiased nonlinear average age-appropriate brain templates from birth to adulthood. *NeuroImage* **47**, S102 (2009).
- Ciric, R. *et al.* TemplateFlow: FAIR-sharing of multi-scale, multi-species brain models. *Nat Methods* **19**, 1568–1571 (2022).
- Jenkinson, M., Bannister, P., Brady, M. & Smith, S. Improved Optimization for the Robust and Accurate Linear Registration and Motion Correction of Brain Images. *NeuroImage* **17**, 825–841 (2002).
- Greve, D. N. & Fischl, B. Accurate and robust brain image alignment using boundary-based registration. *NeuroImage* **48**, 63–72 (2009).
- Jenkinson, M. & Smith, S. A global optimisation method for robust affine registration of brain images. *Medical Image Analysis* **5**, 143–156 (2001).
- Power, J. D. *et al.* Methods to detect, characterize, and remove motion artifact in resting state fMRI. *NeuroImage* **84**, 320–341 (2014).

40. Behzadi, Y., Restom, K., Liau, J. & Liu, T. T. A component based noise correction method (CompCor) for BOLD and perfusion based fMRI. *NeuroImage* **37**, 90–101 (2007).
41. Esteban, O. *et al.* MRIQC: Advancing the automatic prediction of image quality in MRI from unseen sites. *PLOS ONE* **12**, e0184661 (2017).
42. Wang, S. *et al.* An fMRI Dataset for Concept Representation with Semantic Feature Annotations. *Sci Data* **9**, 721 (2022).
43. Smith, D. V., Ludwig, R. M., Dennison, J. B., Reeck, C. & Fareri, D. S. An fMRI Dataset on Social Reward Processing and Decision Making in Younger and Older Adults. *Sci Data* **11**, 158 (2024).
44. Etzel, J. A. *et al.* The Dual Mechanisms of Cognitive Control dataset, a theoretically-guided within-subject task fMRI battery. *Sci Data* **9**, 114 (2022).
45. Mumford, J. A., Turner, B. O., Ashby, F. G. & Poldrack, R. A. Deconvolving BOLD activation in event-related designs for multivoxel pattern classification analyses. *NeuroImage* **59**, 2636–2643 (2012).
46. Mumford, J. A., Davis, T. & Poldrack, R. A. The impact of study design on pattern estimation for single-trial multivariate pattern analysis. *NeuroImage* **103**, 130–138 (2014).
47. Abdulrahman, H. & Henson, R. N. Effect of trial-to-trial variability on optimal event-related fMRI design: Implications for Beta-series correlation and multi-voxel pattern analysis. *NeuroImage* **125**, 756–766 (2016).
48. Woolrich, M. W., Ripley, B. D., Brady, M. & Smith, S. M. Temporal Autocorrelation in Univariate Linear Modeling of FMRI Data. *NeuroImage* **14**, 1370–1386 (2001).
49. Smith, S. M. *et al.* Advances in functional and structural MR image analysis and implementation as FSL. *NeuroImage* **23**, S208–S219 (2004).
50. Smith, S. M. Fast robust automated brain extraction. *Human Brain Mapping* **17**, 143–155 (2002).
51. Guo, T., Liu, X., Chen, M., Fu, Y. & Guo, T. An fMRI dataset for investigating language control and cognitive control in bilinguals. *OpenNeuro* <https://doi.org/10.18112/openneuro.ds005455.v1.1.5> (2024).
52. Magnotta, V. A., Friedman, L. & FIRST BIRN. Measurement of Signal-to-Noise and Contrast-to-Noise in the fBIRN Multicenter Imaging Study. *J Digit Imaging* **19**, 140–147 (2006).
53. Forman, S. D. *et al.* Improved Assessment of Significant Activation in Functional Magnetic Resonance Imaging (fMRI): Use of a Cluster-Size Threshold. *Magnetic Resonance in Medicine* **33**, 636–647 (1995).
54. Birn, R. M. Quality control procedures and metrics for resting-state functional MRI. *Front. Neuroimaging* **2**, (2023).
55. Krüger, G. & Glover, G. H. Physiological noise in oxygenation-sensitive magnetic resonance imaging. *Magnetic Resonance in Medicine* **46**, 631–637 (2001).
56. Fransson, P. & Marrelec, G. The precuneus/posterior cingulate cortex plays a pivotal role in the default mode network: Evidence from a partial correlation network analysis. *NeuroImage* **42**, 1178–1184 (2008).
57. Anticevic, A. *et al.* The role of default network deactivation in cognition and disease. *Trends in Cognitive Sciences* **16**, 584–592 (2012).
58. Hickok, G. *et al.* A functional magnetic resonance imaging study of the role of left posterior superior temporal gyrus in speech production: implications for the explanation of conduction aphasia. *Neuroscience Letters* **287**, 156–160 (2000).
59. Graves, W. W., Grabowski, T. J., Mehta, S. & Gupta, P. The Left Posterior Superior Temporal Gyrus Participates Specifically in Accessing Lexical Phonology. *Journal of Cognitive Neuroscience* **20**, 1698–1710 (2008).

## Acknowledgements

This study was supported by the National Natural Science Foundation of China (31871097) to Taomei Guo. We would like to express our deepest gratitude to the participants who volunteered to participate in this study. We also thank Linyan Liu for assistance with pipeline programming and data processing.

## Author contributions

Tingting Guo: Formal analysis, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. Xuedi Liu: Writing – review & editing. Mo Chen: Investigation, Data curation. Yongben Fu: Investigation, Data curation. Taomei Guo: Conceptualization, Funding acquisition, Investigation, Project administration, Resources, Supervision, Validation, Writing – review & editing.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41597-025-05245-9>.

**Correspondence** and requests for materials should be addressed to T.G.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025