



OPEN

DATA DESCRIPTOR

Cognitive tasks, anatomical MRI, and functional MRI data evaluating the construct of self-regulation

Patrick G. Bissett¹✉, Ian W. Eisenberg², Sunjae Shim¹, Jaime Ali H. Rios¹, Henry M. Jones³, McKenzie P. Hagen⁴, A. Zeynep Enkavi⁵, Jamie K. Li¹, Jeanette A. Mumford¹, David P. MacKinnon⁶, Lisa A. Marsch⁷ & Russell A. Poldrack¹

We describe the following shared data from $N = 103$ healthy adults who completed a broad set of cognitive tasks, surveys, and neuroimaging measurements to examine the construct of self-regulation. The neuroimaging acquisition involved task-based fMRI, resting state fMRI, and structural MRI. Each subject completed the following ten tasks in the scanner across two 90-minute scanning sessions: attention network test (ANT), cued task switching, Columbia card task, dot pattern expectancy (DPX), delay discounting, simple and motor selective stop signal, Stroop, a towers task, and a set of survey questions. The dataset is shared openly through the OpenNeuro project, and the dataset is formatted according to the Brain Imaging Data Structure (BIDS) standard.

Background & Summary

We report a dataset acquired as part of an effort to understand the construct of *self-regulation*, which refers to the processes or abilities that are used to serve long-term goals. Self-regulation has been shown to relate to a variety of real-world outcomes, including economic choices, health outcomes, and academic achievement^{1–3}. We operationalized self-regulation as a large heterogeneity of constituent processes that may be interrelated, including attention, set shifting, decision making, temporal discounting, response inhibition, and planning^{4,5}. As a first step in this project, we created a cognitive *ontology* of self-regulation, or an explicit specification of the entities and relationships composing a domain, which is detailed in previous work^{4,5} (data available at⁶).

The dataset detailed in the present data descriptor was acquired to build on the cognitive ontology to measure the neural underpinnings of self-regulation. We measured self-regulation with a suite of converging modalities including structural MRI and functional MRI (fMRI) during tasks, rest and survey responses. The sample included 103 subjects who each completed two 90 minutes scans with nine putative self-regulatory tasks during fMRI: attention network test⁷ (ANT), cued task switching^{8,9}, Columbia card task¹⁰, dot pattern expectancy¹¹ (DPX), delay discounting¹², simple stop signal¹³, motor selective stop signal¹⁴, Stroop¹⁵, and a towers task¹⁶. Additionally, subjects completed resting state fMRI and a set of 40 survey responses within the scanner taken from the Brief Self-Control Scale¹⁷, Grit Scale¹⁸, Carstensen Future Time Perspective¹⁹, UPPS + P²⁰, and the Impulsiveness-Venturesomeness scale²¹. Finally, subjects also completed anatomical scans.

To date, this neuroimaging dataset has been used in 1 paper²². Some aspects of participant recruitment, sample characteristics, study procedures, and tasks are described in this previous paper²², which analyzed these data to evaluate RT modeling techniques in task-based fMRI modeling.

Methods

Participants. All research activities were approved by the Stanford University Institutional Review Board (Protocol #39322). All participants provided informed consent for the open publication of the deidentified data.

Prospective participants for the study were recruited from the Stanford campus and surrounding San Francisco Bay Area using several methods including paper flyers, the Stanford Sona recruitment system, local

¹Department of Psychology, Stanford University, Stanford, USA. ²Credo AI, Credo, USA. ³Department of Psychology, University of Chicago, Chicago, USA. ⁴Department of Psychology, University of Washington, Washington, USA. ⁵Division of the Humanities and Social Sciences, California Institute of Technology, Pasadena, USA. ⁶Department of Psychology, Arizona State University, Los Angeles, USA. ⁷Center for Technology and Behavioral Health, Geisel School of Medicine, Dartmouth College, Stanford, USA. ✉e-mail: pbissett@stanford.edu

newspapers ads, the Poldrack Lab website, online resources such as Craigslist, and through email listservs maintained by the Stanford Psychology Department. All recruited participants met the following criteria: have a minimum 8th grade education, speak English fluently, right-handed, have normal or corrected to normal vision and no color-blindness, are between 18–50 years old, have no current diabetes diagnosis, have no history of head trauma with loss of consciousness, cerebrovascular accident, seizures, neurosurgical intervention, stroke, or brain tumor, have no current major psychiatric disorders (including schizophrenia and bipolar disorder) or substance dependence, are not currently using any medication for psychiatric reasons, are not currently pregnant, and have no other contraindications to MRI.

A total of 113 participants were recruited for the study. 3 participants were dropped during their first scan session due to complications in the scanner leaving 110 participants. An additional 7 subjects did not complete their second scan session, leaving 103 subjects who completed the entire study. The sample and its demographics are described in more detail within the Technical Validation Section below.

A shorter overview of the sample was described in a previous manuscript²².

Procedure. A shorter overview of the study procedure was described in a previous manuscript²².

Scanner details and image extraction. MRI data were acquired on a GE Discovery MR750 3 T scanner using a Nova medical 32-channel head coil, located at the Stanford Center for Cognitive and Neurobiological Imaging (CNI).

Fieldmap. A fieldmap scan was acquired with a dual-echo spiral sequence with the following parameters: TR = 700 ms, TE1 = 4.545 ms, TE2 = 9.1 ms, flip angle = 56, 1.72 * 1.72 * 4 mm voxels. The magnitude image (magnitude.nii.gz) included in the dataset reflects the image from the first TE and the fieldmap image (fieldmap.nii.gz) shows the off-resonance frequency at each voxel in units of Hz.

fMRI. fMRI scans were acquired with single-echo multi-band echo-planar imaging (EPI) sequence. The following parameters were used for data acquisition: TR = 680 ms, multiband factor = 8, echo time = 30 ms, flip angle = 53 degrees, field of view = 220 mm, 2.2 * 2.2 * 2.2 isotropic voxels with 64 slices. The multiband protocol source is GE and the encoding direction state is j.

Anatomical scans. Sagittal 3D T1-weighted scans were acquired using GE's BRAVO sequence with TR = 7.24 ms, echo time = 2.784 ms, inversion time = 450 ms, flip angle = 12 degrees, matrix = 256 * 256, 0.9 * 0.9 * 0.9 mm voxels. T2-weighted scans were acquired with GE's Cube sequence with TR = 2500 ms, echo time = 95.455 ms, flip angle = 90 degrees, matrix = 320 * 320, 0.5 * 0.5 * 0.8 mm voxels.

Cardiac and respiratory recordings. Cardiac and respiratory data was collected with a pulse oximeter and respiration belt. The respiration belt, pulse oximeter, and emergency squeeze ball are all integrated parts of the GE scanner. We have included these raw files in the data directory within the subject and session matched with each run.

The in-scanner behavioral data was obtained using the fORP 932 response box alongside the PYKA 5 button box, held in participants' right hand, both manufactured by Current Designs.

Participant procedure. The study consisted of 3 parts which included 2 MRI sessions and an at-home self paced online battery hosted on the Experiment Factory²³ (<https://www.expfactory.org>), which is a framework for deploying experiments in the browser. Each MRI session consisted of a 1 hour practice and setup period prior to a 1.5 hours long scanning session for a total of 2.5 hours. All MRI data was obtained at Stanford's CNI.

Before the first MRI session, participants provided informed consent and completed a demographic survey that asked for age, sex, race, and ethnicity. This information is shared in the participants.tsv within the dataset shared.

During both scan sessions, subjects practiced the five tasks that they will do in that session in an observation room on a laptop prior to being scanned. During this practice session, a researcher read the instructions to the subject and the subject had the opportunity to ask any clarifying questions. Sessions were broken up into two groups. Group 1 included the following tasks: Stop Signal task, cued task switching task (two by two), Columbia Card Task (CCTHot), ANT, and Ward and Allport Tower task (WATT). Group 2 included the following tasks: Motor Selective Stop Signal task, DPX task, Survey Medley, Delayed-Discounting task (DDT), and the Stroop task. Within-session task orderings were counterbalanced across subjects by employing 4 different orderings. Separately for the two stop signal tasks, stop signal delay (SSD) was tracked during practice using a staircase algorithm and the final SSD was subsequently used as the starting SSD during the in-scanner version of the tasks. For the two stop signal tasks, response mapping for each shape were randomly decided for each participant during practice and were carried over into the in-scanner version. For the DPX task the valid cue and probe pairing was randomly chosen and used within the practice then carried over to the testing phase. For the task switching task, the color choice order and magnitude choice order were randomly chosen during practice and carried over to the testing phase. All other tasks did not counterbalance response mapping across subjects. The tasks were practiced in the same order that they would be completed within the scanner. The tasks varied in the number of trials and performance requirements detailed below.

The following four tasks contained no performance criteria to proceed onto the scanner version outside of displaying a verbal understanding of the task instructions. The Survey Medley practice included 2 survey questions, one likert scale and another yes/no question, WATT included 4 practice trials, CCTHot included 7 practice trials, and the DDT included 4 practice trials.

The following tasks contained task feedback as well as performance criteria for participants to meet prior to continuing to the scanner version. The ANT practice consisted of 1 block of 16 trials that participants repeated up to a maximum of 3 blocks or until they met the performance threshold of greater than 75% accuracy. For ANT and all other tasks with a similar maximum of 3 practice blocks, if they did not meet the performance criteria after the 3 practice blocks they were still advanced to the main task. The DPX practice consisted of 1 block of 30 trials that participants repeated for a maximum of 3 blocks or until they met the performance threshold of greater than 75% accuracy. The Stop Signal practice consisted of 1 block of 20 trials where the participant practiced only the go trials and repeated this block up to a maximum of 3 times or until they achieved greater than 75% accuracy. After meeting the set performance criteria on go trials, participants completed a set of 2 blocks of 30 trials which included both go and stop trials; these blocks had no performance criteria or feedback. The Motor Selective Stop Signal task practice was the exact same as the Stop Signal task practice. The Stroop task practice consisted of 1 block of 9 trials that participants repeated up to a maximum of 3 blocks or until they achieved greater than 75% accuracy. After meeting the set performance criteria, participants completed 1 block of 20 trials that did not contain any feedback or reminder of instructions/mappings and had no performance criteria to meet. Lastly, The Task Switching task practice consisted of 1 block of 32 trials that participants repeated up to a maximum of 3 blocks or until they achieved greater than 75% accuracy. After meeting the set performance criteria, participants completed 1 block of 16 trials that did not contain any feedback or reminder of instructions/mappings and had no performance criteria to meet.

The time in between practice and testing phases of these tasks was subject to variability depending on the time it took to set up the participant within the scanner and the previous team vacating the scanner, but it roughly averaged around 15 minutes.

Scan sessions consisted of running the following sequences in order: localizer, shim, a single band reference (sbref), rest fMRI, two task fMRI runs, 2nd sbref, three task fMRI runs, and a fieldmap scan. T1-weighted scans were acquired during the first session, and T2-weighted scans were acquired during the second session. If time permitted, additional T1-weighted and T2-weighted scans were acquired during each session. Scan protocols were manually ended based on task run times, leading to variations in scan lengths. All subjects were instructed to refrain from movement during scans. Task instructions were presented again during the fMRI sessions and subjects were also given instructions to respond as quickly and accurately as possible.

Participants were compensated \$20/hr for their participation in the MRI sessions and \$10/hr for participation in the practice prior to scanning. Participants were also compensated \$100 for their completion of a ~10 hour online behavioral battery after the scans. The post-scan battery matched the procedure that each subject completed in Eisenberg *et al.*⁵ and was acquired to allow us to compare the performance in the 10 in-scanner tasks to a broader set of self-regulation tasks and surveys. Quality assurance and processing of these data are ongoing, and once completed we intend to share the data in the OSF repository²⁴ with a matching subject identifier.

For 6 subjects, certain tasks were not run due to the scanner running out of time or other errors (Sub-s061/ses-1: ANT, Sub-s572/ses-1: discountFix, sub-607/ses-2: ANT, sub-609/ses-2: ANT, sub-s623/ses-2: CCTHot, sub-s647/ses-2: discountFix). For 9 task runs, the task continued to run in the background while the scanner protocol ended early, due to running out of time or other errors (sub-s251/ses-2: motorSelectiveStop, sub-s491/ses-2: discountFix, sub-s495/ses-2: discountFix, sub-s555/ses-1: cued task switching, sub-s601/ses-2: cued task switching, sub-s618/ses-1: surveyMedley, sub-s648/ses-2: CCTHot, sub-s649/ses-1: discountFix, sub-s649/ses-2: CCTHot). For 1 task run, the scanner was left running for longer than the task duration (sub-s648/ses-2: ANT). It is recommended to cut off the trailing end of this scan for analysis. For 2 subjects, the scanner had to be restarted, and subjects waited in the scanner in between scan runs (sub-s618/ses-1, sub-s619/ses-1). 4 subjects came out of the scanner to go to the bathroom or adjust their position before reentering the scanner and resuming the rest of the scans for the same session (sub-s613/ses-2, sub-623/ses-1, sub-623/ses-2, sub-s640/ses-1, sub-s649/ses-2). These instances of breaks between tasks are noted in the BIDS dataset in scans.tsv files under each session directory. For 1 subject, one of the task behavioral data was not saved correctly due to the scanner task failing (sub-s499/ses-02, missing discountFix behavior file). For 2 subjects, all behavioral data from the second session was lost (sub-s592/ses-2, sub-s640/ses-2).

Resting state fMRI. During both scan sessions, participants completed about 8 minutes of resting state fMRI scans before starting the task fMRI scans. During these scans, participants were instructed to keep their eyes open and look at the fixation cross on the screen.

fMRI tasks. The following tasks (Fig. 1) were selected from a larger set of 36 tasks that were administered in Eisenberg *et al.*^{4,5}. We used a genetic algorithm to select a subset of tasks that best reconstructed the entire behavioral results across subjects in an independent dataset (see⁴ for additional details). In-scanner tasks were run using a local instantiation of Experiment Factory²³, with modifications to include in-scanner practice trials and accept sync pulses from the scanner. The task started after receiving 16 scan pulses to account for calibration scans at the beginning of each sequence. Behavioral data outputs were stored locally and converted to event files using scripts detailed in Data Records.

All tasks made use of variable inter-trial intervals (ITIs, though see ANT) and fMRI optimized trial orderings generated through the Neurodesign python package²⁵. Neurodesign offers a four-objective optimization that allows the user to weigh four different measures of efficiency: estimation efficiency, detection power, confound efficiency, and frequency efficiency. Estimation efficiency is focused on modeling the time course of the signal, specifically using a finite impulse response function model, which was not our planned modeling approach. Detection power focuses on the power of a model using canonical hemodynamic response function-convolved regressors, which was the planned modeling approach. The other two efficiencies focus on the predictability

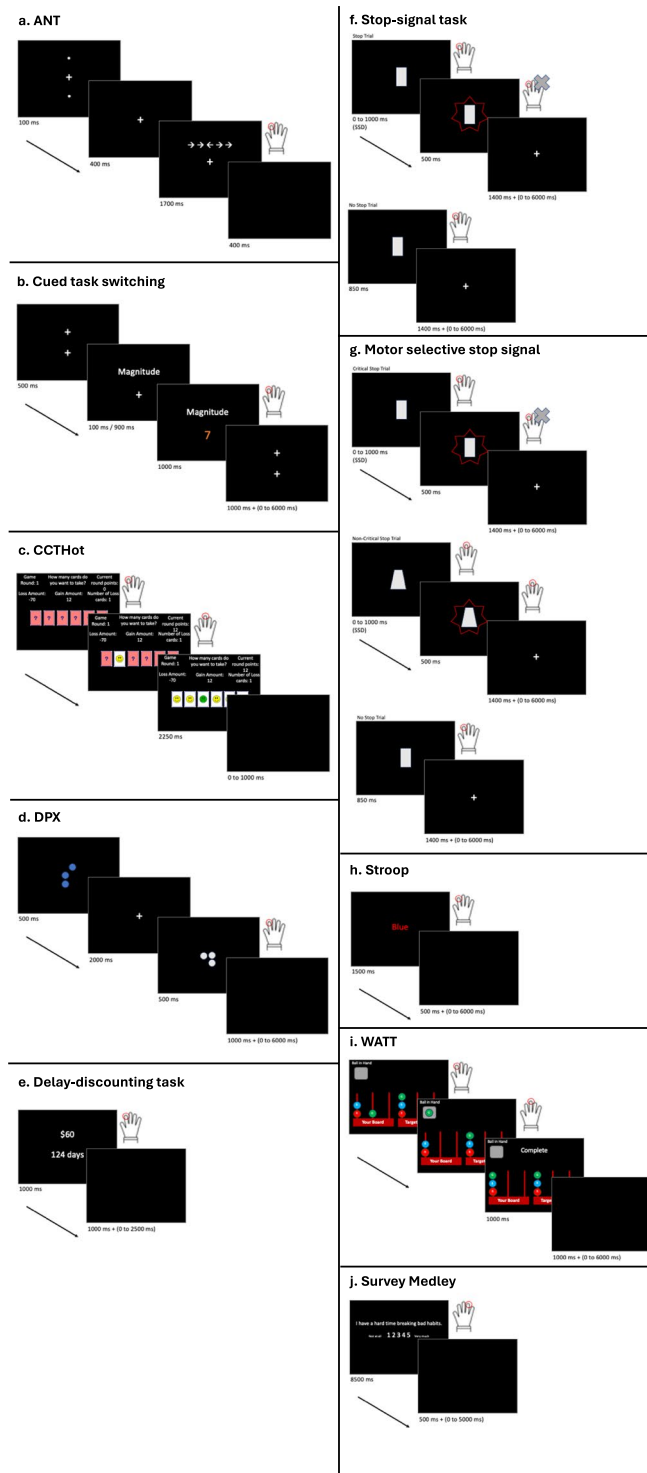


Fig. 1 Schematic of tasks. **(a)** ANT. **(b)** Cued task switching. **(c)** CCTHot. **(d)** DPX. **(e)** Discount delaying task. **(f)** Stop-signal task. Schematic of a typical stop-signal task. The go stimulus stays on screen for 850 ms. If the stop signal delay (SSD) is less than 350 ms, only the go stimulus will be on screen until the SSD, then both go stimulus and stop signal will be on screen for 500 ms, followed by just the go stimulus on screen for 850ms-500ms-SSD. If the SSD is greater than 350 ms, the stop signal was on screen, along with the fixation cross, even after the go stimulus disappeared. **(g)** Motor selective stop signal. Schematic of a typical motor selective stop signal task. The go stimulus stays on screen for 850 ms. If the stop signal delay (SSD) is less than 350 ms, only the go stimulus will be on screen until the SSD, then both go stimulus and stop signal will be on screen for 500 ms, followed by just the go stimulus on screen for 850ms-500ms-SSD. If the SSD is greater than 350 ms, the stop signal was on screen, along with the fixation cross, even after the go stimulus disappeared. Participants had to make a response even if a stop-signal was presented if the go stimulus was a non-critical shape. **(h)** Stroop. **(i)** WATT. **(j)** Survey medley.

| Task | Session Group | Randomization Type | # of Blocks | Total Trials | Average Run Duration |
|-----------------------------|---------------|--------------------|-------------|--------------|----------------------|
| ANT | Group 1 | Neurodesign | 2 | 128 Trials | 347 seconds |
| Two By Two | Group 1 | Neurodesign | 3 | 240 Trials | 677 seconds |
| CCTHot | Group 1 | Neurodesign | 2 | 87 Rounds | 592 seconds |
| DPX | Group 2 | Neurodesign | 4 | 160 Trials | 729 seconds |
| Delay-discounting | Group 2 | Neurodesign | 2 | 120 Trials | 614 seconds |
| Stop Signal | Group 1 | Neurodesign | 3 | 125 Trials | 324 seconds |
| Motor Selective Stop Signal | Group 2 | Neurodesign | 5 | 250 Trials | 651 seconds |
| Stroop | Group 2 | Neurodesign | 2 | 96 Trials | 209 seconds |
| WATT | Group 1 | Neurodesign | 3 | 48 Rounds | 503 seconds |
| Survey Medley | Group 2 | Neurodesign | 1 | 40 Questions | 395 seconds |

Table 1. Description of task parameters. Session Group refers to the task grouping each task is associated with. Scans were split into 2 sessions for each participant and task groupings were counterbalanced across participants. Neurodesign was used to generate ITTs and optimize trial orderings to increase power for specific task contrasts.

of trials due to trial order (confound efficiency) and whether the desired trial type frequency is achieved (frequency efficiency). We chose values of 0, 0.1, 0.5 and 0.4 for estimation efficiency, detection power, confound efficiency and frequency efficiency, respectively, given our primary goals were to minimize predictability while retaining the desired trial frequencies with a secondary goal of maximizing detection power. Four sets of ITTs and trial orderings were then generated by rotating both the ITTs and trial ordering of the originally chosen Neurodesign set, each starting at a different time point within the chosen design of each task to match the four different task orders used across subjects for counterbalancing purposes. Each task used Neurodesign to power for the specific contrasts outlined below. Due to Neurodesign spanning several generated task designs in order to improve power, the condition probabilities can vary from the target probabilities. Unless otherwise specified, Neurodesign was used to generate both trial orderings and ITTs, with the distribution that the ITTs were sampled from and conditions being powered for being listed below for each task.

The tasks parameters are described below and in Table 1. Shorter descriptions of each task are described in²².

Attention network test (ANT). The ANT⁷ is designed to engage networks involved in three putative attentional functions: alerting, orienting, and executive control. Each trial begins with a central fixation point and a cue presented simultaneously. The cue can either be a double cue (one presented at the top and bottom of the screen, which are the two possible target locations) or a spatial cue (deterministically indicates the upcoming target location as top or bottom). A set of five arrow stimuli follows the cue at the top or bottom of the screen, and the subject must indicate via button press the direction of the arrow presented in the center of the five stimuli. The center and four flanking arrows can either point the same direction (congruent trial), or the center and the four flanking arrows can point in opposite directions (incongruent trial). Arrows can only point right or left. Responses are made with the right hand index finger if the center arrow is pointing left, and middle finger if the center arrow is pointing right.

The task consisted of 128 total trials separated into 2 blocks. Each trial consisted of a cue (100 ms), fixation cross (400 ms), probe (1700 ms) plus an added response window with a blank screen (400 ms), and ITI (400 ms) where a fixation is shown on the screen. The onset timing of the probe is specified in the event files included in the dataset. Responses were accepted during the probe and ITI period for each trial. Due to a coding error, this was the one task that did not include a variable inter-trial interval.

Conditions included 16 possible combinations of cue (double, spatial) * probe location (up, down) * probe direction (left, right) * condition (congruent, incongruent), each equally likely to occur. Neurodesign²⁵ was used to generate trial orders equally powering for cue, probe location, probe direction, and condition.

Cued task-switching task (two by two). We used a modified cued task switching task that paired 2 different tasks with 2 different cues^{8,9}. Each trial begins with a fixation followed by a central cue instructing the subject on whether to make a magnitude or a color judgment to the subsequent probe number. The cue is either “high-low”, “magnitude”, “orange-blue”, or “color”, with the former two cues indicating one task (judge whether the subsequent probe number is greater than or less than 5), and the latter two cues indicating the other task (to judge whether the subsequent probe number is orange or blue). The cue was presented 100 ms or 900 ms before the target onset (with equal probabilities) and stayed on the screen throughout the probe phase. The probe was a single colored letter between 1–9 excluding 5. Responses are made with the index and middle finger of the participant’s right hand and the magnitude and color key mappings were randomly chosen during each subject’s practice session and used in the scanner.

The task consisted of 240 trials comprising 3 blocks of 80 trials. Each trial consisted of a fixation (500 ms), cue (100, 900 ms), probe (1000 ms), and a blank-screen ITI (1000 ms + variable ITI, see below for details). The onset timing of the probe is specified in the event files included in the dataset. Responses were accepted in the probe and ITI periods.

Conditions included cue switch (25% of trials), cue stay (25% of trials), and task switch (50% of trials). Cue switch trials are when the current trial has a different cue than the immediately preceding trial, but both indicate the same task (e.g., current trial cue is “magnitude” and previous cue was “high-low”). Cue stay trials are when the current trial has the exact same cue as the immediately preceding trial (e.g., current trial cue is “magnitude” and the immediately preceding trial cue was also “magnitude”). Task switch trials are when the current trial has a cue that indicates a different task than the task that from the immediately preceding trial (e.g., current trial cue is “magnitude” and the previous cue was “color”). The variable portion of the ITI was sampled from an exponential distribution using a min of 0, mean of 0.25, and a max of 6 seconds, and conditions were powered for task switch - (cue switch + cue stay) and cue switch - cue stay contrasts.

Columbia card task (CCTHot). The CCTHot task¹⁰ has subjects play a card game with the goal of collecting as many points as possible by flipping over gain cards and avoiding flipping over loss cards. Each round presents a set of 6, 8, 9, 10, 12, 15, or 16 cards face down with the following information about the cards: The number of loss cards (1, 2, 3, or 5 possible), the amount they will lose if they turn over a loss card (ranging from -100 to -5), and the amount they earn with a gain card (ranging from 1 to 30). The subject can choose whether to flip over a card at random by pressing the index finger key or “cash out” and move to the next trial by pressing the middle finger key. A subject may flip over zero or as many cards as they choose, but the trial ends when they flip over a loss card. After each choice is made, the subject is shown which card they flipped over. For each round, they will be awarded the net number of points from the gain and loss cards they flipped over.

The task consists of 87 rounds split into 2 blocks of 44 and 43 rounds. Each trial is subject paced. After each round there is a feedback phase of 2250 ms with an added blank-screen variable ITI (minimum of 0, mean of 0.2 s, max of 1 s). The scan was stopped after 12 minutes, so given the variable trial duration this resulted in a different number of trials per subject (M of completed trials across participants = 81.02, SD = 12.75, range = 32–87).

Dot pattern expectancy task (DPX). In the DPX task¹¹, subjects see a series of cue-probe stimulus pairs and make a speeded response after the probe. Each stimulus is made with a pattern of dots. There are 6 possible cues, and 6 possible probes. One cue is the “target cue” (i.e. ‘A’) and one probe is the “target probe” (i.e. ‘X’). Subjects must press one key if the target cue is followed by the target probe (i.e. ‘AX’), and another key for any other cue-probe pairing. The target cue and target probe was randomly chosen during practice and input for the test phase. Responses are made with the right hand index finger for AX pairs and middle finger for any other pair.

The task consists of 4 blocks of 40 trials for a total of 160 trials. Each trial consisted of a cue (500 ms), fixation (2000ms), probe (500 ms), and a blank inter-trial interval (1000 ms + variable ITI). Probe responses could be made during the 500 ms probe or during the inter-trial interval. The onset timing of the cue and probe are specified in the event files included in the dataset. There are four trial conditions, AX, AY, BX, and BY, with probabilities 0.55, 0.15, 0.15, 0.15, respectively. AX trials are a target cue preceding the target probe, AY trials are a target cue preceding a non-target probe, BX trials are a non-target cue preceding a target probe, and BY trials are a non-target cue preceding a non-target probe. ITIs were sampled from a truncated exponential distribution with a minimum of 0, a mean of 0.4 s, and a max of 6 s. We powered for AY-BY and BX-BY contrasts.

Delay-discounting task (DDT). In the delayed-discounting task¹², subjects are presented with a dollar amount and a number of days. They are asked to choose between receiving \$20 today or the larger amount of money at the presented delay in days. The amount of money (\$22 - \$85) and number of days (19 days - 180 days) varies across trials. Subjects were told that their performance will impact the bonus they receive. Subjects made a right hand middle finger button press for the “smaller sooner” reward of \$20 and an index finger button press for the “larger later” presented reward. The task consists of 2 blocks of 60 trials for a total of 120 trials. Each Trial consists of a probe (4000 ms) and a blank screen ITI (500 ms + variable ITI). Responses were accepted to both the probe and the ITI.

ITIs were sampled from a truncated exponential distribution with a minimum of 0, mean of 0.5 s, and max of 2.5 s.

Stop-signal task. In the stop-signal task¹³, subjects are presented with one of 4 “go” shapes (moon, oval, trapezoid, rectangle) and are instructed to respond to two shapes with one key press and the other two shapes with another key press. On a subset of trials, a star appears around the shape after a delay (the stop-signal delay, SSD). This indicates to the participant that they should try not to make any response on that trial. If a subject is unable to withhold their response on a stop-signal trial this is categorized as a stop failure, and if the subject makes no response then this is categorized as a stop success. During both the practice and main task, the SSD was tracked using a “1 up 1 down” staircase algorithm²⁶ in which the SSD increased by 50 ms after each successful stop signal trial and decreased by 50 ms after each failed stop signal trial. The final SSD in practice was used as the starting SSD during the in-scanner version of the tasks.

There are a total of 125 trials, grouped into 3 blocks (41 trials in block 1 and 42 trials in each of blocks 2 and 3), with 60% of the trials being go trials and 40% being stop trials. The stop signal duration was 500 ms and the initial SSD during practice was 250 ms. Each trial consisted of a go stim (850 ms) and an inter-trial interval that included a central fixation cross (1400 ms + variable ITI). ITIs were sampled from an exponential distribution using a min of 0, mean of 0.225, and a max of 6 seconds. We powered for the stop - go contrast. Go responses were made with the index and middle finger of the participant’s right hand and the key mapping was randomly chosen during each subject’s practice session and used in the scanner.

Motor selective stop-signal task. The motor selective stop-signal task¹⁴ is identical to the previously described stop-signal task, with the following exceptions. Instead of instructing subjects to stop whenever a stop signal occurs, subjects are instructed to only stop if a stop signal occurs and they were going to make one of their two responses (the “critical” response), but not if a stop signal occurs and they were going to make the other response (the “non-critical” response). SSD was adjusted only on critical stop trials, and the SSD on noncritical stop trials was yoked to the SSD on the critical stop trials. Responses are made with the index and middle finger of the participant’s right hand and the key mapping is randomly chosen during each subject’s practice session and used in the scanner. Critical stop trial stimuli are also randomly chosen during the subject’s practice session and used in the scanner.

The task consists of 5 blocks of 50 trials for a total of 250 trials. 60% of trials were go trials, 20% critical stop trials, and 20% noncritical stop trials. We powered for the following contrasts: Noncritical stop - noncritical no-signal and critical stop - noncritical stop.

Stroop task. In the Stroop task¹⁵, subjects are presented with a color word (e.g., “red”) written in ink that either matches the word (congruent, e.g., “red” in red) or does not (incongruent, e.g., “red” in blue). Subjects are instructed to quickly and accurately respond via keypress to the ink color of the word is. The color of the word and the written word could both be red, blue or green. The participants were asked to respond with index, middle, and ring finger for each of these colors.

The task consists of 2 blocks of 48 trials for a total of 96 trials. Each trial consists of a probe (1500 ms) and an inter-trial interval that includes a central fixation cross (500 ms + variable ITI). There were 48 congruent and 48 incongruent trials. ITIs were sampled from an exponential distribution using a min of 0, mean of 0.2, and a max of 6 seconds. We powered for the incongruent - congruent contrast.

Ward and Allport tower task (WATT). In the WATT¹⁶ subjects are presented with stacked balls in a specific initial configuration on three pegs, and they are asked to move the stacked balls to a different target configuration using the least amount of moves.

The task consists of 3 blocks of 16 rounds for a total of 48 rounds. These moves are not timed and the round ends when the subject moves the ball to match the target configuration shown on the screen. All trials were designed to be completed in 3 steps, where a step is defined by taking a ball off the board then placing it on a different peg. After the participant correctly solves each round there is a feedback phase that consists of feedback shown for 1000 ms and a blank screen (1000 ms + variable ITI). Some subjects did not complete the full 48 rounds, as there was a max scan time of 10 minutes and the task was ended after 10 minutes even if all trials were not completed (M of completed rounds across participants = 45.35, $SD = 4.99$, range = 28–48). Participants practiced with trials in which the target configuration had all three balls positioned on the same peg (Unambiguous). The main test phase consisted of trials where the target configuration had one ball on one peg and two balls on another peg (Partially Ambiguous). Trials included conditions in which a ball had to be moved out of the way to reach the target position (with intermediate step) and in which all balls could directly be moved to the target position (without intermediate step). Both practice and test phase were scanned.

ITIs were sampled from an exponential distribution using a min of 0, mean of 0.9, and a max of 6 seconds. We powered for the intermediate vs. without intermediate step contrast.

Survey medley. In addition to the above 9 traditional cognitive tasks, subjects also completed 40 survey questions that were answered within the scanner. We scanned the following three complete surveys: the Grit Scale¹⁸ (8 questions), Brief Self-Control Scale¹⁷ (13 questions), and the Carstensen Future Time Perspective¹⁹ (10 questions, Carstensen). We also selected a subset of items from the Impulsiveness-Venturesomeness²¹ (3 questions) and UPPS + P impulsivity²⁰ (6 questions) questionnaires. Our previous work showed minimal relation between self-regulatory tasks and self-regulatory surveys, and showed that self-regulatory surveys were more predictive of real-world outcomes than self-regulatory tasks⁵. We scanned a subset of the surveys from Eisenberg *et al.*⁵ in order to explore the neural underpinning of self-regulatory surveys.

Each trial presented the survey question for 8500 ms and was followed by a blank screen for 500 ms plus a variable ITI (minimum of 0, mean of 1 s, max of 5 s).

Data Records

This data descriptor provides a description of this neuroimaging dataset, which is openly shared on OpenNeuro²⁷ under accession number ds004636. The shared data are organized according to the Brain Imaging Data Structure (BIDS), which is a data organization structure designed to encourage FAIR sharing and reuse of neuroimaging data²⁸. Processed behavioral data are included in the OpenNeuro repository for each task as events.tsv files. Raw behavioral data of fMRI tasks is available on OpenScience Framework²⁴. This includes sidecar JSON files with detailed explanation of MRI scan protocol parameters as well as task parameters. Functional scans are labeled as follows:

- ANT: Attention network test
- twoByTwo: Cued task switching task
- CCTHot: Columbia card task
- DPX: Dot pattern expectancy
- discountFix: delayed discounting
- stopSignal: Stop-signal task
- motorSelectiveStop: Motor selective stop-signal task
- Stroop: Stroop task

| Task | Condition | Reaction Time | Accuracy |
|---|------------------------|---------------|---------------|
| ANT (N = 87) | Spatial Incongruent | 0.605 (0.108) | 0.977 (0.041) |
| | Double Incongruent | 0.628 (0.092) | 0.973 (0.041) |
| | Spatial Congruent | 0.569 (0.092) | 0.990 (0.019) |
| | Double Congruent | 0.592 (0.091) | 0.986 (0.038) |
| Two By Two (Task Switching Task) (N = 88) | Task Switch | 0.764 (0.123) | 0.892 (0.078) |
| | Cue Switch | 0.735 (0.125) | 0.926 (0.077) |
| | Cue Stay | 0.699 (0.112) | 0.941 (0.059) |
| DPX (N = 87) | AX | 0.522 (0.112) | 0.963 (0.058) |
| | AY | 0.645 (0.114) | 0.906 (0.099) |
| | BX | 0.433 (0.138) | 0.947 (0.082) |
| | BY | 0.452 (0.137) | 0.970 (0.053) |
| Simple Stop Signal (N = 88) | Go | 0.655 (0.140) | 0.945 (0.059) |
| | Stop Failure | 0.587 (0.110) | 0.964 (0.050) |
| Motor Selective Stop Signal (N = 81) | Non-Critical Signal | 0.613 (0.110) | 0.938 (0.089) |
| | Non-Critical No Signal | 0.579 (0.100) | 0.960 (0.063) |
| | Critical Stop Failure | 0.633 (0.119) | 0.900 (0.100) |
| | Critical Go | 0.703 (0.154) | 0.932 (0.080) |
| Stroop (N = 90) | Congruent | 0.632 (0.098) | 0.970 (0.057) |
| | Incongruent | 0.732 (0.140) | 0.940 (0.064) |

Table 2. Post Exclusion Reaction time and Accuracy Metrics. Reaction times are in means across subjects in seconds. Standard deviations are presented in parentheses.

| Task | Measure | Average |
|--------------------------------------|---|---------------|
| Columbia Card Task (CCTHot) (N = 90) | Rewarded Trials Proportion | 0.537 (0.070) |
| | Average number of cards pulled per round | 3.490 (1.122) |
| | Average number of loss cards pulled per round | 0.170 (0.033) |
| Delay-discounting Task (N = 88) | Larger Later Choice Proportion | 0.527 (0.293) |
| | Discount Rate | 0.019 (0.033) |
| Simple Stop Signal (N = 88) | Stop Signal Delay | 0.363 (0.137) |
| | Stop Success Rate | 0.520 (0.037) |
| | Stop Signal Reaction Time | 0.270 (0.051) |
| | Go Omission Rate | 0.013 (0.036) |
| Motor Selective Stop Signal (N = 81) | Stop Signal Delay | 0.342 (0.161) |
| | Stop Success Rate | 0.520 (0.053) |
| | Stop Signal Reaction Time | 0.337 (0.081) |
| | Non-Critical Signal Omission Rate | 0.038 (0.067) |
| | Non-Critical No Signal Omission Rate | 0.015 (0.059) |
| WATT (N = 89) | Critical Go Omission Rate | 0.018 (0.058) |
| | Partially ambiguous with intermediate step - extra moves | 0.204 (0.332) |
| | Partially ambiguous without intermediate step - extra moves | 0.099 (0.178) |

Table 3. Post Exclusion Measures other than RT and accuracy. Standard deviations are presented in parentheses.

- WATT3: towers task
- surveyMedley: 40 survey questions
- rest: Resting state (eyes open)

The participants.tsv file contains age and sex of each participant, and suggested_exclusions.csv contains notes on suggested exclusions based on behavior and neuroimaging data quality assurance (described below).

Cardiac and respiratory recordings are included as gzip compressed.tsv files for each functional run. The cardiac data files are named in the following way: sub-`{subject_number}_ses-{session_number}_task-{task_name}_run-{run_number}_recording-cardiac_physio.tsv.gz. The respiratory data files are named in the following way: sub-{subject_number}_ses-{session_number}_task-{task_name}_run-{run_number}_recording-respiratory_physio.tsv.gz. A JSON file with relevant metadata is included in the main directory. As noted in the JSON file, cardiac and respiratory data were recorded starting from 30 seconds before the functional time series started.`

| | |
|--|-----|
| Total Subjects | 103 |
| Subjects recommended for complete exclusion | 12 |
| Subjects recommended for partial exclusion | 23 |
| Subjects with no recommended exclusion | 68 |
| ANT suggested inclusions | 87 |
| Two By Two suggested inclusions | 88 |
| CCTHot suggested inclusions | 90 |
| DPX suggested inclusions | 87 |
| Delay-discounting Task suggested inclusions | 88 |
| Stop Signal suggested inclusions | 88 |
| Motor Selective Stop Signal suggested inclusions | 81 |
| Stroop suggested inclusions | 90 |
| WATT suggested inclusions | 89 |
| Survey Medley suggested inclusions | 91 |

Table 4. Suggested exclusions/inclusions by task.

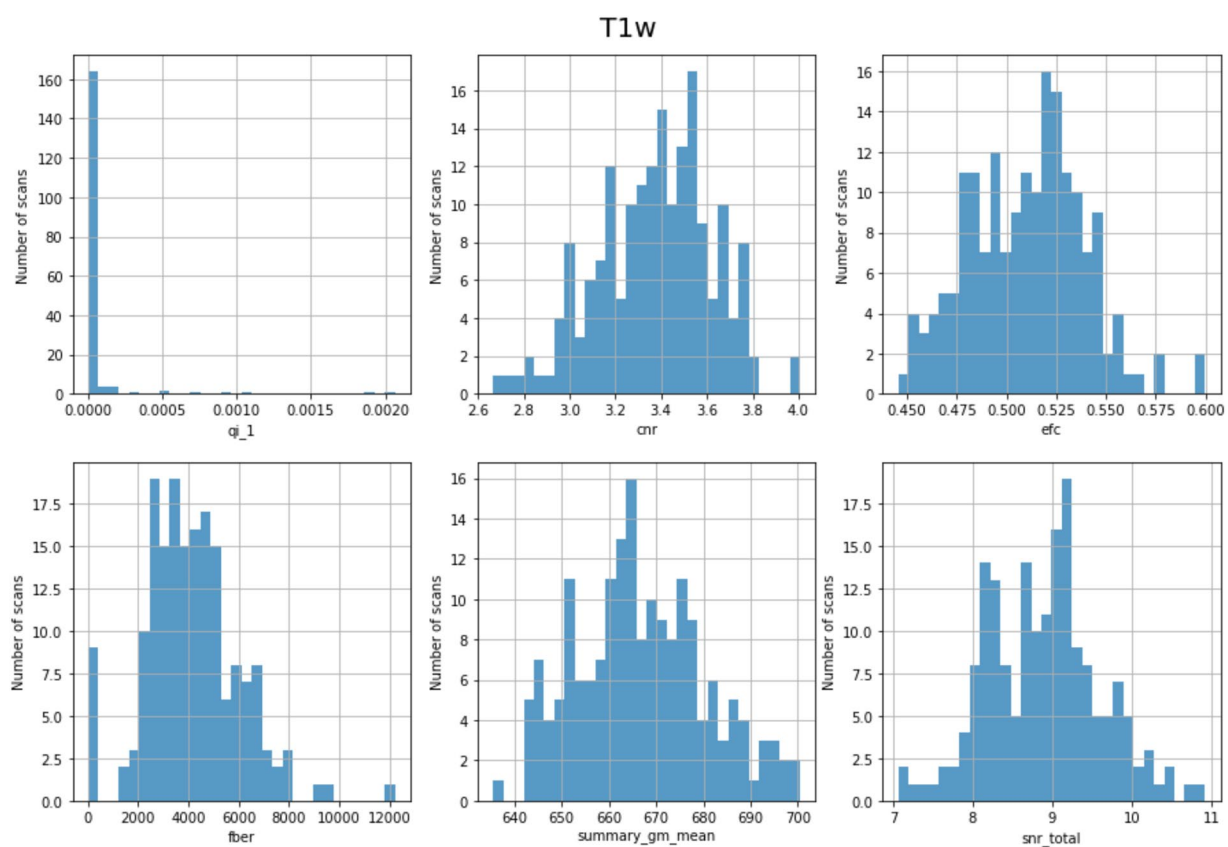


Fig. 2 Distribution of QC measures of T1-weighted data. See Table 5 for metric descriptions.

Derivatives. All anatomical and functional scan images were run through MRIQC version 22.0.6²⁹ (RRID:SCR_022942). T1-weighted and T2-weighted images were defaced using pydeface version 2.0.0³⁰ prior to being put through MRIQC. MRIQC reports are shared as Hypertext Markup Language (html) files and image quality metrics (IQM) metadata are shared as JSON files in the derivatives/mriqc directory within the root BIDS directory. BIDS derivatives are included as a directory on OpenNeuro ds004636.

Behavioral data. Raw task behavior during scans were collected using Experiment Factory²³(<https://www.expfactory.org>) based on jsPsych⁵¹. These data include the in-scanner practice portion of the task that was not scanned, if it was included in task structure. All raw and cleaned behavioral data are shared in an OSF repository²⁴. Detailed information about each column in the raw behavioral outputs are explained in Aim1 DataDescriptor Expfactory Output Column Descriptions.pdf within the OSF repository under Reference Files. Raw data were cleaned and processed to create event files using python script named process_data.py, which takes

| Image Quality Metric (Label) | Description | Range across all T1 images (Median) |
|---|--|-------------------------------------|
| Artifact detection (qi_1) | Detects artifacts from motion, blurring or ghosting ³⁴ . qi_1 values are the percent of artifact voxels, calculated as the proportion of artifact corrupted voxels normalized by the number of background voxels. Lower values are better. | 0.0–0.002 (0.0) |
| Contrast-to-noise ratio (cnr) | Measure of contrast between white and gray matter ³⁵ . Calculated as $\frac{\text{Mean}_{\text{gray matter}} - \text{Mean}_{\text{white matter}}}{\text{Standard Deviation}_{\text{air}}}$. Higher values indicate a larger difference between white and gray matter signal intensities. Higher values are better. | 2.666–4.006 (3.397) |
| Entropy focus criterion (efc) | Proxy for ghosting or blurriness caused by head motion ³² . Lower values are better. | 0.446–0.600 (0.513) |
| Foreground-background energy ratio (fber) | Variance of voxels inside the brain divided by variance of voxels outside of the brain ³⁶ . Higher values are better. | -1.0–12216.989 (3956.549) |
| Gray matter signal (summary_gm_mean) | Average signal of gray matter. | 635.546–700.432 (665.494) |
| Signal-to-noise ratio (snr_total) | Relative measure of background noise. Calculated as $\frac{\mu_F}{\sigma_F \sqrt{n/(n-1)}}$ where μ_F is the mean intensity of tissues, σ_F is the standard deviation of tissues and n is the number of voxels in the tissue mask. Higher values are better. | 7.062–10.923 (8.922) |

Table 5. Description of QC measures of T1-weighted anatomical data.

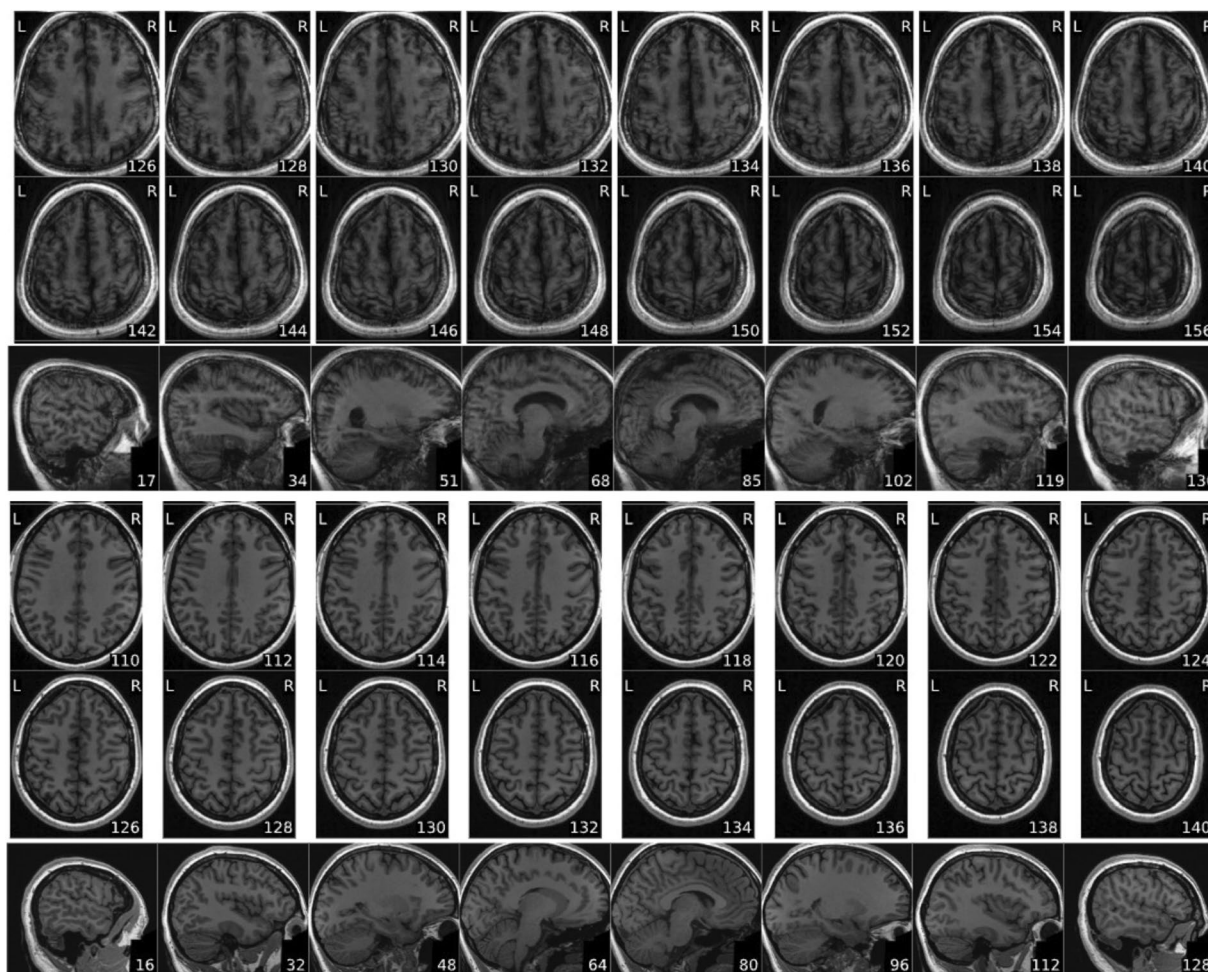


Fig. 3 T1w image by slices of scans with lowest (top) and highest (bottom) signal-to-noise ratio value from MRIQC reports. top) MRIQC report for sub-s499_ses-1_T1w.nii.gz. This subject was suggested as an exclusion in suggested_exclusions.csv. bottom) MRIQC report for sub-s358_ses-1_T1w.nii.gz.

raw_behavioral_files from OSF repo as input and produces cleaned_behavioral_files in OSF repo and event_files in OpenNeuro dataset³². Detailed sidcar files for each task's event file can be found in the BIDS dataset along with the neuroimaging data on OpenNeuro following BIDS specifications. Data are organized by subject number and task name. We chose to separate out the code (in Zenodo) and behavioral data (in OSF) from the neuroimaging data (in OpenNeuro) to make each modality easier to navigate in a more tailored repository.

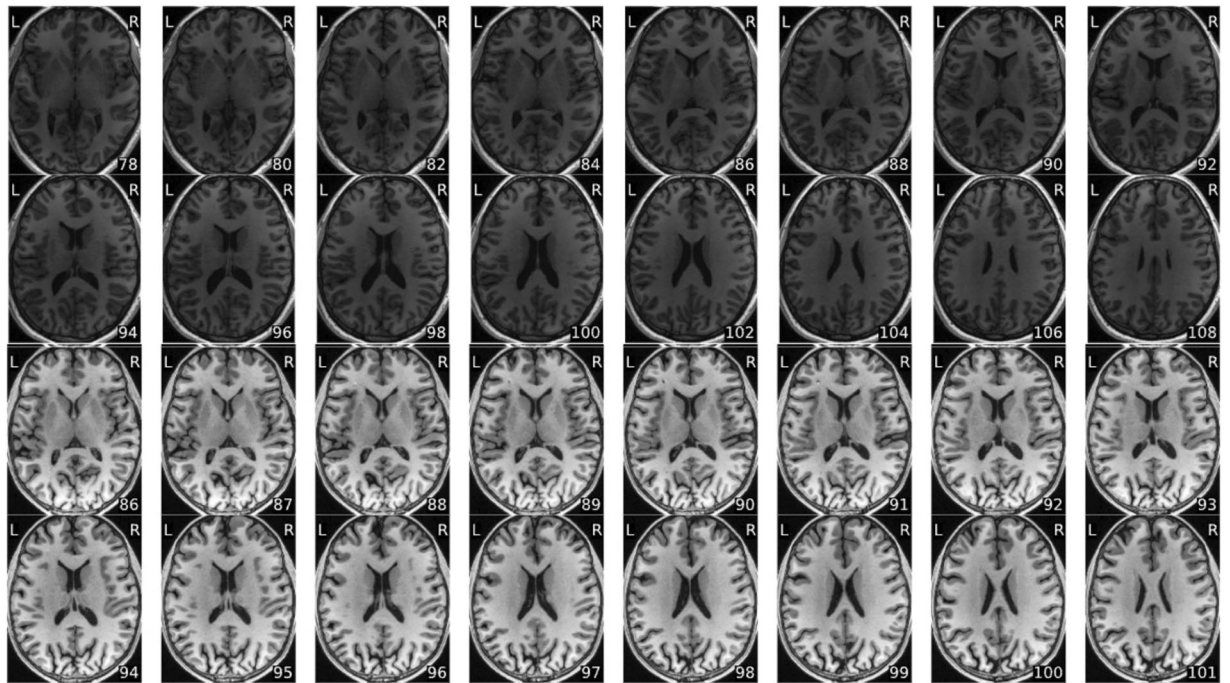


Fig. 4 T1w image by slices of scans with lowest (top) and highest (bottom) foreground-background energy ratio value from MRIQC reports. top) MRIQC report for sub-s172_ses-2_T1w.nii.gz bottom) MRIQC report for sub-s130_ses-1_T1w.nii.gz.

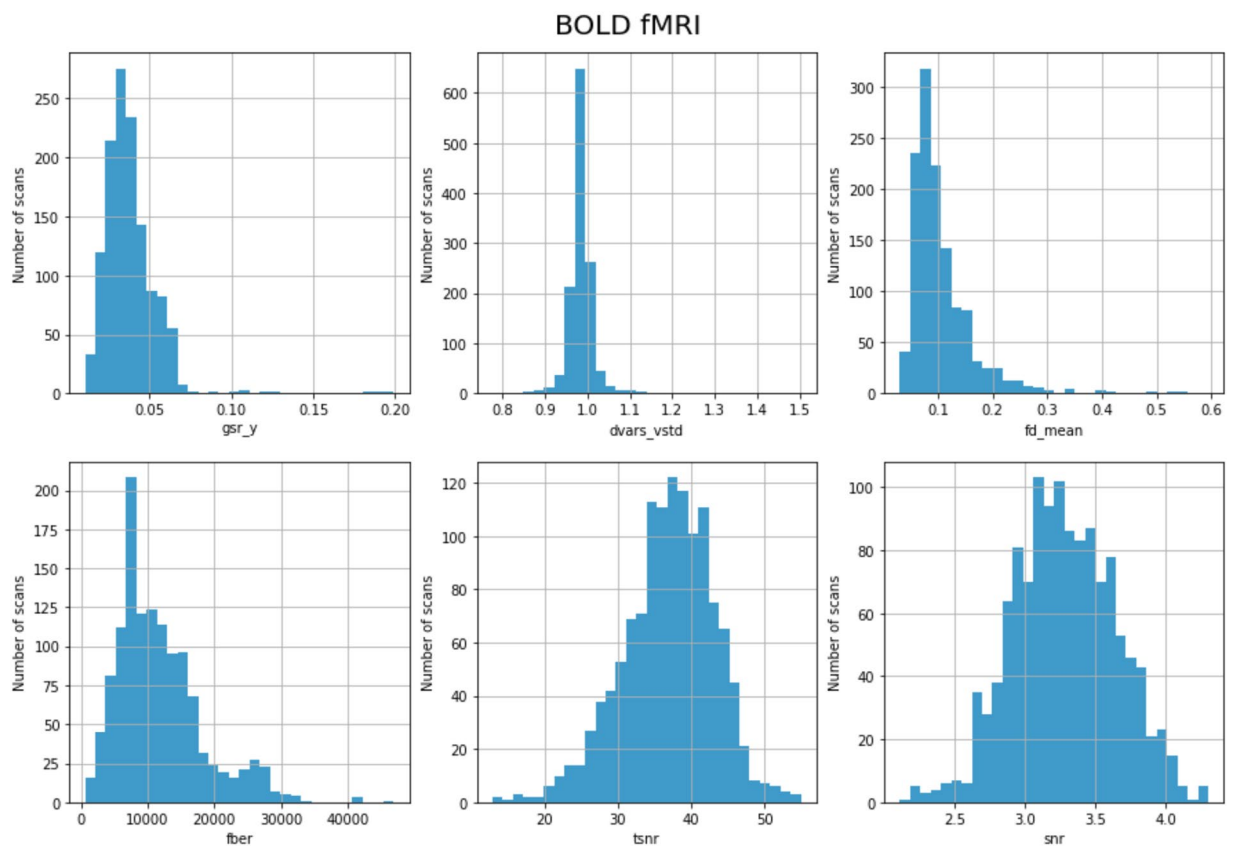


Fig. 5 Distribution of QC measures for BOLD functional scans. See Table 6 for metric descriptions.

| Image Quality Metric (Label) | Description | Range across all functional scans (Median) |
|---|---|--|
| Ghost to Signal Ratio (gsr_y) | Measure of signal intensity in areas along encoding direction, subtract the signal intensity in areas not along the encoding direction, and then divide the resultant value by the signal intensity in the brain ³⁷ . Lower values are better. | 0.011–0.199 (0.036) |
| DVARS (dvars_vstd) | Measure of signal changes from volume to volume. Lower values are better. | 0.777–1.503 (0.984) |
| Framewise displacement (fd_mean) | Measure of motion throughout the scan measured in millimeters. Lower values are better. | 0.030–0.595 (0.089) |
| Foreground-background energy ratio (fber) | Average value of energy within the head compared to regions outside of the head ³⁶ . Higher values are better. | 574.653–46811.020 (10483.188) |
| Temporal SNR (tsnr) | Relative measure of background noise across time course. tsnr value is the median value of $\frac{\text{average BOLD signal(across time)}}{\text{temporal standard deviation map}}$ that is calculated for each timepoint of the image. Higher values are better. | 12.884–54.966 (37.324) |
| Signal-to-noise ratio (snr) | Relative measure of background noise. Calculated as $\frac{\mu_F}{\sigma_F \sqrt{n/(n-1)}}$ where μ_F is the mean intensity of tissues, σ_F is the standard deviation of tissues and n is the number of voxels in the tissue mask. Higher values are better. | 2.111–4.304 (3.266) |

Table 6. Description of QC measures of BOLD data.

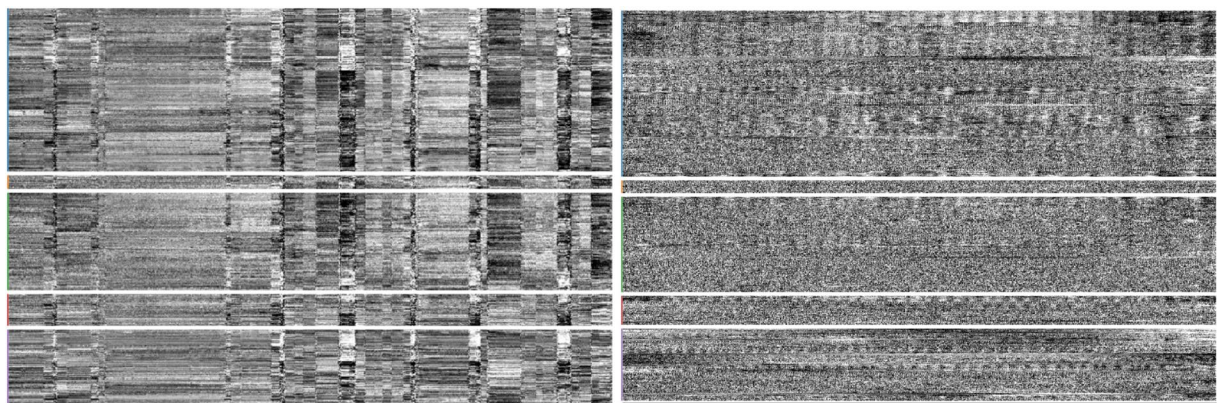


Fig. 6 Carpet plot of scan with highest (left) and lowest (right) average framewise displacement value from MRIQC reports. left) MRIQC report for sub-s605_ses-1_task-WATT3_run-1_bold.nii.gz. Note large disconnects in carpet plot. We do not suggest any motion-based exclusion, and we encourage users to implement their own motion-based exclusion. right) MRIQC report for sub-s130_ses-1_task-rest_run-1_bold.nii.gz.

Technical Validation

All collected data are included in the current data release. Quality assurance is described in detail in the immediately following sections, and we recommend that certain subjects and scans be excluded, but we recognize that exclusion decisions can depend on the goals of each project, so we have decided to include all available data in the data release so that data users can make exclusion decisions that align with their use cases.

Behavioral quality control. We completed the following behavioral quality control, and we have provided annotations of these quality control recommendations in suggested_exclusions.csv on OpenNeuro.

Falling outside the normal range on any of the following criteria was sufficient for that subject's data to be recommended for exclusion. In the stop-signal tasks, stop success rate being below 0.25 or above 0.75, in line with consensus recommendations³³. In all tasks, having an omission rate of over 0.5. Subjects missing more than half the tasks had their entire dataset (all tasks) recommended for exclusion. We ran a binomial test on noncritical signal trials in motor selective stopping to ensure that subjects were making responses on more than half of trials. If they failed this binomial test, they were recommended for exclusion as it indicated they were treating the task as a simple stop signal task. Last, we looked to see if any subjects omitted a majority of their responses at the beginning or end of any scans, but this did not identify any additional subjects.

We also evaluated descriptive statistics for each subject on each task, including their accuracy, RT, and other key dependent variables (e.g., range of SSDs for the stop-signal task). This resulted in the recommended exclusion of 11 tasks across 6 subjects, mostly for accuracy that was at or below chance in at least one condition. Details of each of these 11 decisions can be found in the subjective_rating_descriptions.json within the BIDS directory. We have also included summary statistics for each task across the 91 subject post-exclusion dataset in Tables 2 and 3. Summary statistics of the full dataset (N = 103) can be found in our OSF repo.

As a result of this quality assurance effort, we recommend that 12 subjects be completely excluded because performance in at least 5 of their 9 speeded tasks (i.e., all but the survey medley) were deemed unsatisfactory. Additionally, we identified 23 subjects who had satisfactory behavior on at least 5 of their 9 tasks but we recommend excluding their data for 1 to 4 of their tasks based upon performance. Therefore, of the 103 participants

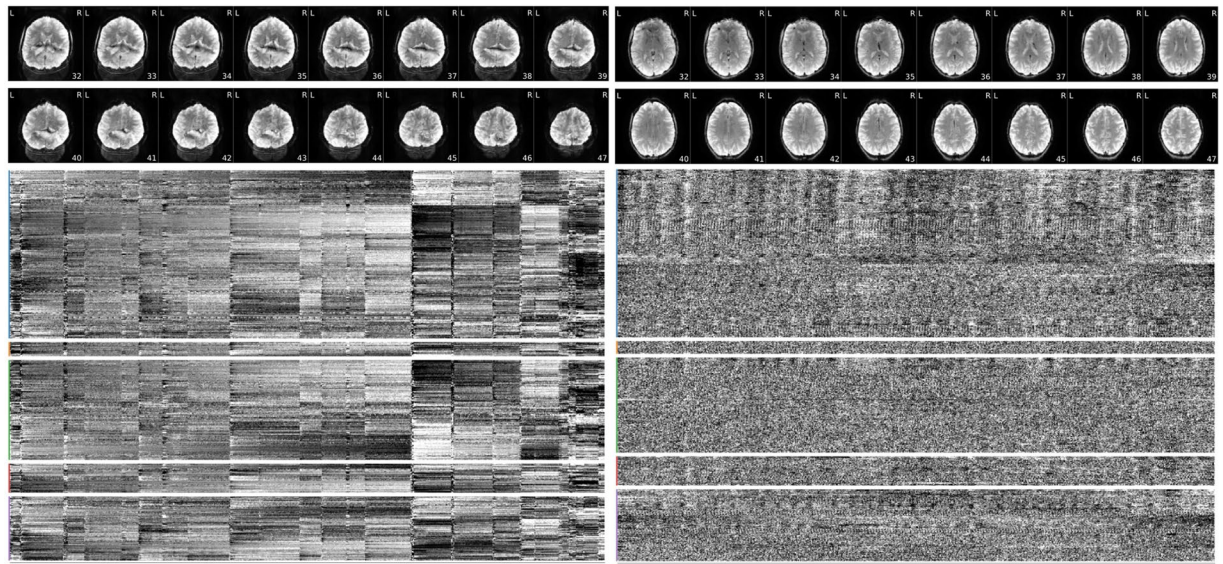


Fig. 7 BOLD average image by slice and carpet plot of scan with lowest (left) and highest (right) temporal signal-to-noise ratio value from MRIQC reports. left) MRIQC report for sub-s499_ses-1_task-WATT3_run-1_bold.nii.gz. Note severe ghosting in BOLD average images and multiple spikes in carpet plot suggesting high motion. This subject was suggested as an exclusion in suggested_exclusions.csv due to missing more than half of task data and having suboptimal functional scan MRIQC results. right) MRIQC report for sub-s130_ses-1_task-surveyMedley_run-1_bold.nii.gz.

in the dataset, 68 are judged to be both complete and have satisfactory behavioral data on all tasks. Suggested exclusions by task are listed in Table 4.

The full dataset ($N = 103$) has the following demographics: 65% Female, 35% Male, 43% White, 34% Asian, 10% More than one race, 9% Black or African American, 2% Unknown, 2% Native Hawaiian or Pacific Islander and a mean age of 24. We suggest that 12 participants should be excluded. The sample after this $N = 12$ exclusion ($N = 91$; 63% Female, 37% Male; Age 18–41, mean: 25) consists of the following demographic distribution: 47% White, 32% Asian, 9% More than one race, 9% Black or African American, 2% Unknown, and 1% Native Hawaiian or Pacific Islander.

We suggested that 35 participants have all or part of their data excluded (12 full, 23 partial). This subsample ($N = 35$; 57% Female, 43% Male; Age: 18–45, mean: 26) had the following demographics: 41% Asian, 41% White, 6% More Than One Race, 6% Black or African American, 3% Unknown, and 3% Native Hawaiian or Other Pacific Islander. In general, our included and excluded samples tended to have more females than males, younger adults, and predominantly White or Asian individuals.

Neuroimaging analyses and quality control. Imaging files were converted from DICOM data to Neuroimaging Informatics Technology Initiative (NIfTI) format using NIMSDa (<https://github.com/cni/nimstda>). T1-weighted anatomical images and BOLD contrast (fMRI) images were processed with MRIQC²⁹ (RRID:SCR_022942) to extract the various image quality metrics listed below. MRIQC image quality metrics outputs are included in OpenNeuro derivatives. Scripts used to generate the attached figures are included in the following repository³².

Anatomical scans. Anatomical T1w images were defaced using pydeface version 2.0.0³⁰ and the defaced images were used to calculate the following measures for quality control using MRIQC²⁹ (RRID:SCR_022942). See Fig. 2 for T1w quality assurance metrics and Table 5 for an explanation of the quality control metrics and range across all T1w images. We did not exclude any additional subjects for anatomical scan quality, but some of the lowest quality anatomical scans were already excluded based upon other criteria (e.g., the participant in the top panel of Fig. 3).

We've included some examples of MRIQC visual reports from scans with highest and lowest values for some of the quality control metrics. See Figs. 3 and 4 for examples of minimum and maximum value scans of signal-to-noise ratio and foreground-background energy ratio. This is intended to exemplify some of the key metrics and show the range over which they vary in this dataset.

Functional BOLD scans. Quality control metrics were computed using MRIQC for each fMRI scan. MRIQC reports were visually inspected for visible artifacts such as aliasing, distortion, and ghosting. Specific scans with MRIQC reports that did not pass inspection are noted in suggested_exclusions.csv within the dataset. Survey medley data have not been closely evaluated and therefore we do not make any suggested exclusions for this task. See Fig. 5 for distributions of quality control measures across all functional scans (task and rest) and Table 6 for an explanation of the quality control metrics and range across all functional scans. We computed the

average of each quality control measure for each functional scan and found them to be highly similar across tasks (gsr_y ($M = 0.038$, $SD = 0.001$, range = 0.037–0.040), dvars_vstd ($M = 0.986$, $SD = 0.012$, range = 0.967–1.010), fd_mean ($M = 0.108$, $SD = 0.008$, range = 0.092–0.120), fber ($M = 11672.360$, $SD = 379.258$, range = 11149.039–12441.484), tsnr ($M = 36.842$, $SD = 2.646$, range = 33.108–42.351), snr ($M = 3.281$, $SD = 0.020$, range = 3.247–3.310). Therefore, in Fig. 5, we present all task and rest data aggregated.

We've included examples of MRIQC visual reports from scans with highest and lowest values for some of the quality control metrics. See Figs. 6 and 7 for examples of lowest and highest value scans of framewise displacement and temporal signal-to-noise ratio, respectively.

Code availability

All scripts used to create event files and to create figures in this manuscript are available here³².

Received: 27 October 2023; Accepted: 11 July 2024;

Published online: 20 July 2024

References

- Mischel, W., Shoda, Y. & Rodriguez, M. L. Delay of Gratification in Children. *Science* **244**, 933–938 (1989).
- Moffitt, T. E. *et al.* A gradient of childhood self-control predicts health, wealth, and public safety. *Proc. Natl. Acad. Sci.* **108**, 2693–2698 (2011).
- Duckworth, A. L. & Seligman, M. E. P. Self-Discipline Outdoes IQ in Predicting Academic Performance of Adolescents. *Psychol. Sci.* **16**, 939–944 (2005).
- Eisenberg, I. W. *et al.* Applying novel technologies and methods to inform the ontology of self-regulation. *Behav. Res. Ther.* **101**, 46–57 (2018).
- Eisenberg, I. W. *et al.* Uncovering the structure of self-regulation through data-driven ontology discovery. *Nat. Commun.* **10**, 2319 (2019).
- Eisenberg, I. *et al.* Uncovering the structure of self-regulation through data-driven ontology discovery. *OSF* <https://doi.org/10.17605/OSF.IO/ZK6W9> (2023, September 6).
- Fan, J., Mccandliss, B., Fossella, J., Flombaum, J. & Posner, M. The activation of attentional networks. *NeuroImage* **26**, 471–479 (2005).
- Logan, G. D. & Bundesen, C. Clever homunculus: Is there an endogenous act of control in the explicit task-cuing procedure? *J. Exp. Psychol. Hum. Percept. Perform.* **29**, 575–599 (2003).
- Mayr, U. & Kliegl, R. Differential effects of cue changes and task changes on task-set selection costs. *J. Exp. Psychol. Learn. Mem. Cogn.* **29**, 362–372 (2003).
- Figner, B., Mackinlay, R. J., Wilkening, F. & Weber, E. U. Affective and Deliberative Processes in Risky Choice: Age Differences in Risk Taking in the Columbia Card Task. *J. Exp. Psychol. Learn. Mem. Cogn.* **35**, 709–730 (2009).
- Otto, A. R., Skatova, A., Madlon-Kay, S. & Daw, N. D. Cognitive Control Predicts Use of Model-based Reinforcement Learning. *J. Cogn. Neurosci.* **27**, 319–333 (2015).
- Kable, J. W. *et al.* No Effect of Commercial Cognitive Training on Brain Activity, Choice Behavior, or Cognitive Performance. *J. Neurosci.* **37**, 7390–7402 (2017).
- Logan, G. D. & Cowan, W. B. On the Ability to Inhibit Thought and Action: A Theory of an Act of Control. *Psychol. Rev.* **91**, 295–327 (1984).
- Aron, A. R., Behrens, T. E., Smith, S., Frank, M. J. & Poldrack, R. A. Triangulating a Cognitive Control Network Using Diffusion-Weighted Magnetic Resonance Imaging (MRI) and Functional MRI. *J. Neurosci.* **27**, 3743–3752 (2007).
- Stroop, J. R. Studies of interference in serial verbal reactions. *J. Exp. Psychol.* **18**, 643–662.
- Kaller, C. P., Rahm, B., Spreer, J., Weiller, C. & Unterrainer, J. M. Dissociable Contributions of Left and Right Dorsolateral Prefrontal Cortex in Planning. *Cereb. Cortex* **21**, 307–317 (2011).
- Roth, R. M., Isquith, P. K., & Gioia, G. A. BRIEF-A: Behavior rating inventory of executive function - adult version.
- Duckworth, A. L. & Quinn, P. D. Development and Validation of the Short Grit Scale (Grit-S). *J. Pers. Assess.* **91**, 166–174 (2009).
- Carstensen, L. L. & Lang, F. R. Future time perspective scale. Unpublished Manuscript
- Cyders, M. A. *et al.* Integration of impulsivity and positive mood to predict risky behavior: Development and validation of a measure of positive urgency. *Psychol. Assess.* **19**, 107–118 (2007).
- Eysenck, S. B. G., Pearson, P. R., Easting, G. & Allsopp, J. F. Age norms for impulsiveness, venturesomeness and empathy in adults. *Personal. Individ. Differ.* **6**, 613–619 (1985).
- Mumford, J. A. *et al.* The response time paradox in functional magnetic resonance imaging analyses. *Nat. Hum. Behav.* **8**, 349–360 (2023).
- Sochat, V. V. *et al.* The Experiment Factory: Standardizing Behavioral Experiments. *Front. Psychol.* **7**, (2016).
- Bissett, P., Shim, S. & Rios, J. A. H. Cognitive tasks, anatomical MRI, and functional MRI data evaluating the construct of self-regulation. *OSF* <https://doi.org/10.17605/OSF.IO/SJVNW> (2024, May 22).
- Durnez, J., Blair, R. & Poldrack, R. A. *Neurodesign: Optimal Experimental Designs for Task fMRI*. <http://biorxiv.org/lookup/doi/10.1101/119594>, <https://doi.org/10.1101/119594> (2017).
- Levitt, H. Transformed up-down methods for psychoacoustics. *J. Acoust. Soc. Am.* **49**, 467–477 (1971).
- Bissett, P. G. *et al.* Cognitive tasks, anatomical MRI, and functional MRI data evaluating the construct of self-regulation. *OpenNeuro*. <https://doi.org/10.18112/openneuro.ds004636.v1.0.4> (2024).
- 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).
- Esteban, O. *et al.* MRIQC: Predicting Quality in Manual MRI Assessment Protocols Using No-Reference Image Quality Measures. **18**.
- Gulban, O. F. *et al.* pydeface. *Zenodo* <https://doi.org/10.5281/zenodo.3524401> (2019).
- De Leeuw, J. R., Gilbert, R. A. & Luchterhandt, B. jsPsych: Enabling an Open-Source Collaborative Ecosystem of Behavioral Experiments. *J. Open Source Softw.* **8**, 5351 (2023).
- Shim, S., Rios, J. A. H. & Bissett, P. G. Update scripts for Cognitive tasks, anatomical MRI, and functional MRI data evaluating the construct of self-regulation. *Zenodo*. <https://doi.org/10.5281/zenodo.11211816> (2024).
- Verbruggen, F. *et al.* A consensus guide to capturing the ability to inhibit actions and impulsive behaviors in the stop-signal task. *eLife* **8**, e46423 (2019).
- Mortamet, B. *et al.* Automatic quality assessment in structural brain magnetic resonance imaging: Automatic QA in Structural Brain MRI. *Magn. Reson. Med.* **62**, 365–372 (2009).
- Magnotta, V. A., Friedman, L. & Birn, First Measurement of Signal-to-Noise and Contrast-to-Noise in the fBIRN Multicenter Imaging Study. *J. Digit. Imaging* **19**, 140–147 (2006).

36. Shehzad, Z. *et al.* The preprocessed connectomes project quality assessment protocol - a resource for measuring the quality of MRI data. *Front. Neurosci.* (2019).
37. Giannelli, M., Diciotti, S., Tessa, C. & Mascalchi, M. Characterization of Nyquist ghost in EPI-fMRI acquisition sequences implemented on two clinical 1.5 T MR scanner systems: effect of readout bandwidth and echo spacing. *J. Appl. Clin. Med. Phys.* **11**, 170–180 (2010).

Acknowledgements

This work was supported by the National Institutes of Health (NIH) Science of Behavior Change Common Fund Program through an award administered by the National Institute for Drug Abuse (NIDA) (UH2DA041713; PIs: Marsch, LA & Poldrack, RA).

Author contributions

Eisenberg, I.W., Enkavi, A.Z., Li, J.K., and Bissett, P.G., acquired the data. Bissett, P.G., Eisenberg, I.W., Shim, S., Rios, J.A.H., Jones, H.M., Hagen, M.P., Enkavi, A.Z., Li, J.K., Mumford, J.A., & Poldrack, R.A. quality assured and analyzed the data. Poldrack, R.A., Marsch, L.A., & MacKinnon, D.P., acquired funding for the research and guided the data acquisition.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to P.G.B.

Reprints and permissions information is available at 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 4.0 International License, which permits use, sharing, adaptation, 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 changes were made. 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/4.0/>.

© The Author(s) 2024