A multimodal brain imaging dataset on sleep deprivation in young and old humans

Authors

Gustav Nilsonne[1,2], Sandra Tamm[1,2], Paolo d'Onofrio[1], Hanna Å Thuné[1,3], Johanna Schwarz[1,2], Catharina Lavebratt[4,5], Jia Jia Liu[4,5], Kristoffer NT Månsson[6,7], Tina Sundelin[2,8], John Axelsson[1,2], Predrag Petrovic[2], Peter Fransson[2], Göran Kecklund[1], Håkan Fischer[8], Mats Lekander[1,2], Torbjörn Åkerstedt[1,2]

Affiliations

1. Stockholm University, Stress Research Institute, Stockholm, Sweden
2. Karolinska Institutet, Department of Clinical Neuroscience, Stockholm, Sweden
3. University of Glasgow, Institute of Neuroscience and Psychology, Glasgow, United Kingdom
4. Karolinska Institutet, Department of Molecular Medicine and Surgery, Stockholm, Sweden
5. Karolinska University Hospital, Center for Molecular Medicine, Stockholm, Sweden.
6. Department of Behavioral Sciences and Learning, Linköping University, Linköping, Sweden
7. Department of Adult Psychiatry, PRIMA Barn- och Vuxenpsykiatri, Stockholm, Sweden
8. Stockholm University, Department of Psychology, Stockholm, Sweden

Corresponding author: Gustav Nilsonne (gustav.nilsonne@ki.se)

Abstract

The Stockholm Sleepy Brain Study I is a functional brain imaging study where 48 younger (20-30 years) and 36 older (65-75 years) healthy participants underwent magnetic resonance imaging after normal sleep and partial sleep deprivation in a crossover design. We performed experiments investigating emotional mimicry, empathy for pain, and cognitive reappraisal, as well as resting state functional magnetic resonance imaging (fMRI). We also acquired T1- and T2-weighted structural images and diffusion tensor images. On the night before imaging, participants were monitored with ambulatory polysomnography and were instructed to sleep either as usual or only three hours. Participants came to the scanner the following evening. Besides MRI scanning, participants underwent behavioral tests and contributed blood samples, which have been stored in a biobank and used for DNA analyses. Participants also completed a variety of self-report measures. The resulting multimodal dataset may be useful for hypothesis generation or independent validation of effects of sleep deprivation and aging, as well as investigation of cross-sectional associations between our different outcomes.

Background & Summary

Insufficient sleep is common in everyday life and in psychiatric conditions, but its functional brain imaging correlates are not well known. For this reason, functional magnetic resonance imaging (fMRI) is increasingly used to study effects of sleep deprivation. Examples of domains where effects of sleep deprivation have been investigated include memory, emotion regulation, and intrinsic (resting state) connectivity[1–4].

The Stockholm Sleepy Brain study I aimed to to investigate the effects of partial sleep deprivation (PSD) on resting state brain connectivity, emotional contagion, empathy, and emotional regulation. We included one group of younger (20-30 years) and one group of older (65-75 years) participants. The study had a crossover within-group design. MRI scanning was performed twice with an interval of one month between sessions. The order of partial sleep deprivation or full sleep was counterbalanced. Participants slept in their own homes, a setting which offers high ecological validity compared to a sleep laboratory, and were monitored by ambulatory polysomnography. In the full sleep condition, they were instructed to go to bed and to get up at their usual times. In the sleep deprivation condition, they were instructed to go to bed 3 hours before the time they would usually wake up, and then keep to their usual rise time. Participants could not be blinded to whether or not they were sleep deprived, but the experimenters at the MRI scanner were blinded to the participants' deprivation condition.

Resting state connectivity was investigated using two 8-minute runs with eyes open (figure 1a). Participants were monitored during fMRI scanning by means of eye-tracking to ensure they were awake. Emotional contagion was investigated using pictures of facial emotional expressions presented in blocks, and with electromyography (EMG) of the major zygomatic and superciliary corrugator muscles (figure 1b). Empathy for pain was investigated using pictures of hands stung with needles (figure 1c). Emotional regulation by reappraisal was investigated using picture stimuli with positive and negative valence coupled with instructions to upregulate, downregulate, or maintain their emotional response (figure 1d). All functional imaging experimental paradigms are described in greater detail in the methods section. Structural anatomical brain images and diffusion tensor images were also acquired.

Sleepiness was measured with subjective ratings on the Karolinska Sleepiness Scale[5] and sustained attention was measured with a psychomotor vigilance task. Participants were characterised using biometric measures and a number of self-rating instruments. Blood was collected and used for analysis of leukocyte telomere length and for genetic polymorphisms. DNA and plasma have been stored in a biobank and will be used for further analyses, e.g. of pro-inflammatory cytokines.

Potential reuse scenarios for these data include hypothesis generation and independent validation of research questions and effects related to any of our many outcomes, as well as inclusion in individual participant data meta-analyses.

Methods

Participants

Participants in the younger group were recruited by poster advertising on campus sites in Stockholm and on a website (www.studentkaninen.se). Older participants were recruited by newspaper advertising. Prospective participants were screened for inclusion/exclusion criteria using an online form, and eligibility according to these criteria was confirmed in an interview upon arrival to undergo scanning. Criteria for inclusion were, first, those required to undergo fMRI procedures and use the hand-held response box, namely: no ferromagnetic items in body, not claustrophobic, not pregnant, no refractive error unless it could be corrected using contact lenses to less than five diopters (which could be corrected in the goggles used for

stimulus presentation), not color-blind, and right-handed. In addition, participants were required to be 20-30 or 65-75 years old (inclusive), to have no current or past self-reported psychiatric or neurological illness, including addiction, to not have hypertension or diabetes, to not use psychoactive or immune-modulatory drugs, to not use nicotine every day, and to not have a higher habitual daily caffeine intake than the amount corresponding to four cups of coffee. In the screening data (Data citation 1), it can be seen that one participant reported intake in excess of this criterion, but that participant was nonetheless included and confirmed on site that their habitual daily intake was less than four cups of coffee. A further criterion was to not study, have studied, or be occupied in the fields of psychology, behavioral science, or medicine, including nursing and other allied fields. The reason for this was that participants with a background in psychology might be prone to metacogitate about the experimental paradigm, whereas participants with a background in medicine might have a less strong emotional response to pictures showing needles or injuries, which were used in two of the experiments. The insomnia severity index (ISI) and the Karolinska Sleep Questionnaire (KSQ) were used to exclude participants with insomnia symptoms, out-of-circadian sleep patterns, or snoring/apnea (see below). The Hospital Anxiety and Depression Scale (HADS) was used to exclude participants with depressive symptoms (see below). For practical reasons, participants were also required to understand and speak Swedish fluently and to live in the greater Stockholm area.

Design and procedures

After screening, participants were randomized to undergo sleep deprivation or full sleep first. We performed a block randomization with block sizes of four, and randomization both to sleep condition and to the order of stimulus presentation. The randomization list was generated by one researcher not involved in the MRI scanning (PdO). We visited participants in their home with polysomnography equipment to record their sleep in both conditions. In the next afternoon or evening, the participants reported to the MR center at the Karolinska University Hospital, Stockholm, Sweden.

At the first visit to the MR center, participants were interviewed to confirm that inclusion criteria applied. When participants arrived, they were offered a cheese sandwich and a glass of water. They were given instructions about the experiment and were shown demonstrations of the task-related paradigms. They then completed questionnaires (PANAS, Self-rated health, and Sickness-Q, see below) and a 5-minute psychomotor vigilance test. Participants changed into a hospital gown and entered the MRI scanner, as described below. After scanning, participants performed additional behavioral experiments (see below), blood sampling, and debriefing. Two or three experimenters worked in every session.

Participants were paid 2500 SEK (approx. 280 Euro/360 USD), subject to tax. Those who participated in the Zürich prosocial game (see below) were able to win up to an additional 160 SEK, also subject to tax. Participants were also offered taxi travel to and from the MRI imaging center, in order to avoid traffic accidents following sleep deprivation. Data were collected from 2012-12-04 to 2013-03-27 and from 2013-10-24 to 2014-04-29.

Rating scales

Unless otherwise indicated, rating scales were administered using Google Forms by the participants in their own homes. In order to maintain confidentiality, participants identified themselves only by a code we gave to them when responding to the scales. All scales were used in Swedish translations. The following scales were included:

Insomnia Severity Index (ISI)[6] was used in screening to characterize insomnia severity and exclude participants with high values (≥ 15). We used a validated Swedish translation[7].

Hospital Anxiety and Depression scale (HADS)[8] was used in screening to characterize anxiety and depression severity and exclude participants with high values for the depression subscale (≥ 8). We used a validated Swedish translation[9].

Karolinska Sleep Questionnaire (KSQ) was used in screening to characterize sleep patterns and exclude those with a sleep phase that was too early or too late, defined as usual bedtime before 22:00 or after 01:00. For participants recruited after the summer of 2013, snoring and apnea > 3 times a week were also exclusion criteria. We used the original Swedish version of the scale[10].

Interpersonal Reactivity Index (IRI)[11,12] was used at baseline to characterize trait empathy in participants. The IRI has 4 subscales: empathic concern, perspective taking, personal distress, and fantasy. An attempt to validate the Swedish translation using confirmatory factor analysis on the Swedish translation has shown that the four subscales distribute over two factors, one of which consists of the empathic concern subscale[13].

Toronto Alexithymia Scale-20 (TAS-20)[14] was used at baseline to characterize alexithymia in participants. We used a validated Swedish translation [15].

State-Trait Anxiety Inventory (STAI-T)[16] was used at baseline to characterize anxiety in participants. We used a non-validated Swedish translation available at https://github.com/GNilsonne/Stimulus-presentation-code-Oxazepam-and-emotion/tree/master/Rating-scales.

Big Five Inventory (BFI)[17–19] was used at baseline to characterize participants' personality traits. We used a validated Swedish translation[20].

Perceived Stress Scale-14 (PSS-14)[21] was used at baseline to characterize participants' self-rated psychological stress.

Diurnal Type Scale[22] was used at baseline to characterize participants' self-rated diurnal type.

Questions about resilience to sleep loss[23] were asked at baseline to characterize participants' resilience to sleep loss.

Epworth Sleepiness Scale (ESS)[24] was used at baseline to characterize participants' sleepiness. We used a validated Swedish translation, approved by the Swedish Sleep Research Society, and available at http://www.swedishsleepresearch.com/uploads/Image/ess.pdf.

Positive And Negative Affect Schedule (PANAS)[25] was used at baseline and before scanning to characterize participants' emotional states. We used a validated Swedish translation[26].

Self-rated health was measured at baseline and before scanning using two questions. SRH-5 is a question asking for health in general[27], and we also used a question asking for health right now on a 7-point Likert-type scale.

Sickness-Q is a 10-item scale which we recently developed (Andreasson et al, in revision), and which was used at baseline and before scanning. We used the original Swedish version.

Barratt Impulsiveness Scale-11 (BIS-11)[28] was administered at baseline, and, to a subset of participants, on paper. We used a Swedish version translated and verified by back-translation by Katarina Öberg and Jonas Hallberg, both at Karolinska Institutet, who kindly shared the instrument. The instrument is available at https://github.com/GNilsonne/SleepyBrain-StimulusPresentation/blob/master/RatingScales/Barratt-Impulsiveness-Scale-11.

Karolinska Sleep Diary[29] was administered on paper in order to monitor participants' sleep pattern before and after they participated. Participants were instructed to complete the questionnaire in the morning on 3 days prior to the experiment, on the same day, and on the day after.

Psychopathic Personality Inventory-Revised (PPI-R)[30] was sent to participants before the summer of 2013 by regular mail as a follow-up in order to characterize participants' psychopathic traits. Participants who did not respond were sent a follow-up letter. The first mailing contained a cinema ticket voucher as compensation for completing the form. Participants after the summer of 2013 were given the form at the same time as their sleep diaries. We used a Swedish translation which we have recently validated using data including those from this dataset (Sörman et al, in review). Thus, although the translation has been

validated, this validation is not independent from the present dataset. This form could not be administered online for copyright reasons.

Brown ADD scales (BADD)[31] were administered on paper at the same time as the PPI-R in order to characterize participants' impulsiveness. We used a Swedish translation[32]. This form also could not be administered online for copyright reasons.

Mini Mental Test (MMT)[33,34] was administered on paper to the older participants at their first visit to the MR center in order to verify that they did not have pronounced deficits in cognitive functioning.

Demographic and biometric measures

Participants' sex and age were self-reported at screening and age was recorded in whole years. Height was self-reported at baseline and recorded as whole cm. Education level was self-reported at baseline and coded as "graduated from elementary school", "graduated from secondary school", "currently enrolled in tertiary education", and "graduated from tertiary education". Before entering the MRI scanner room, participants were weighed and weight was recorded in whole kg. To preserve anonymity, BMI but not height and weight are among the variables published openly. To investigate effects of sleep deprivation on physical appearance, participants were photographed. The photographs are not published, since they cannot be de-identified.

Polysomnography

Polysomnography (PSG) recording took place in the homes of the participants using a solid state, portable sleep recorder. Most recordings were made with an Embla system, but a few were made instead with a Vitaport system. Standard electrode (silver/silver chloride) montage for EEG sleep recording was used (C3, C4 referenced to the contralateral mastoid). In addition, two sub-mental electrodes were used for electromyography (EMG) and one electrode at each of the outer canthi of the eyes were used for electrooculography (EOG). Sleep staging, respiratory and arousal analyses were performed according to the classification criteria of the American Academy of Sleep Medicine (AASM)[35]. To adapt to AASM scoring, F4 was interpolated. Here the terminology N2 and N3 is used for sleep stages 1-3. Wake within the total sleep period (WTSP) represents time awake between sleep onset and offset and this value is expressed in percent of the total sleep period (TSP). Shifts from any of the sleep stages to wake were expressed as awakenings per hour.

Experimental paradigms used during fMRI scanning

During scanning, participants were monitored by eye-tracking. In case of eye-closures of more than approx. 5 seconds, the MRI operator spoke a wake-up call through the participant's headphones.

Resting state. Functional connectivity in the resting state was assessed during two eight-minute runs. Participants were instructed to keep their eyes open and look at a white fixation cross presented against a gray background. In the second run, participants rated their sleepiness with the KSS every two minutes.

Emotional mimicry. In a block design, participants were shown happy, angry, or neutral faces from the Karolinska Directed Emotional Faces (KDEF) database[36] (figure 1b). Each block was 20 seconds long and contained 20 faces, each shown for 0.5 seconds. Blocks were arranged in sets of three, beginning and ending with the same emotion, and with neutral in the middle (happy-neutral-happy or angry-neutral-angry). After each set of three, participants were asked to rate how happy and angry they felt on a visual analog scale of 0-100.

Empathy for pain. In an event-related design, participants were shown pictures of hands being stung by needles or poked with Q-tips (figure 1c). The stimulus set was kindly shared by Claus Lamm, and has been shown to be effective in previous work by him[37]. After each picture,

participants were asked to rate their perceived unpleasantness on a visual analog scale of 0-100.

Emotional reappraisal. In an event-related design, participants were instructed to upregulate, maintain or downregulate responses to neutral or negative stimuli (figure 1d). The stimulus pictures were selected from the International Affective Picture System[38]. After each picture, participants were asked to rate how successfully they had regulated their emotional response, on a scale of 1-7. The experimental paradigm was adapted from earlier work by Armita Golkar and colleagues[39].

MRI acquisition

Scanning was performed on a 3T Discovery 750 MRI scanner (General Electric). The full protocol notes have been published together with the MRI data.

T1 anatomical images. T1-weighted images were acquired for normalization of functional images and for morphometric analyses. The following settings were used: field of view 24, slice thickness 1 mm, sagittal orientation, interleaved acquisition bottom to top, covering the whole head. Before publication of images (Data citation 1), the face region was removed in order to preserve participants' anonymity, using a script developed and executed by openfmri.org (https://github.com/poldracklab/pydeface).

Resting state fMRI. Resting state fMRI sequences were acquired using echo-planar imaging (EPI) with field of view 28.8, slice thickness 3 mm, no interslice gap, axial orientation, 49 slices covering the whole brain, interleaved acquisition bottom -> up, TE 30, TR 2.5 s, and flip angle 75.

Task-related fMRI. Sequences for fMRI acquisition were identical for the three task-related experiments (emotional mimicry, empathy for pain, and emotional reappraisal). Before task-related fMRI, we performed a higher order shim procedure with a manually defined region of interest defined during each scan and including approximately the whole frontal, temporal, and parietal lobes. Imaging was optimised for visualising activity in the amygdala. 46 slices with a thickness of 2.3 mm and an interslice gap of 0.1 mm were placed starting at the inferior margin of the pons. Unfortunately, this strategy resulted in missing coverage of both the amygdala and the prefrontal cortex in some participants. These findings will be reported in detail elsewhere (Tamm et al, in preparation). Images were acquired with axial orientation, interleaved bottom to top, TE 34, TR 3 seconds, and flip angle 80°. In most sessions, B0 maps were also acquired corresponding to these fields of view, and were converted to field maps for use in preprocessing of fMRI data in order to reduce spatial distortions.

Diffusion tensor imaging (DTI). DTI images were acquired in order to study the relationship between structural connectivity/white matter integrity and other outcomes. Images were acquired with 45 diffusion directions, TE 80, TR 7 s, field of view 22, slice thickness 2.3, interslice gap 0.1. These data have not been analysed yet.

T2 anatomical images. T2-weighted images were acquired for the purpose of routine evaluation of all participants by the neuroradiology service of the Karolinska University Hospital. The following settings were used: field of view 24, slice thickness 1.2 mm, sagittal orientation, interleaved acquisition bottom -> up, covering the whole head. Before publication of images (Data citation 1), the face region was removed in order to preserve participants' anonymity, as for the T1-weighted anatomical images.

MRI data analysis

In this paper, we report technical validation of resting state experiments.

Preprocessing of resting state fMRI data. 68 participants had complete data from both resting state runs in both sessions. To remove task-related interference, volumes scanned during KSS ratings in the second session (S2) were cut out of the time series, including two volumes before and one after each rating event. Remaining volumes were concatenated and in order to balance data, each series was trimmed down to the lowest number of remaining scans in any

one session, which was 163, corresponding to 7 minutes and 20 seconds. Data were preprocessed in SPM12 using the DPARSF toolbox[40]. Functional images were slice time corrected, realigned, normalized using DARTEL[41], resampled to 3x3x3 mm voxel size, spatially smoothed with a 6x6x6 mm kernel, detrended, frequency filtered (0.01-0.1 Hz), and regressed on nuisance covariates including six motion regressors and gray and white matter signal using DARTEL-obtained segmentation. Participants with more than 40 volumes (25%) with framewise displacement (FD) ≥ 0.5 mm, calculated based on the method of Power et al.[42], were excluded from analysis (n = 15). Remaining volumes with FD ≥ 0.5 mm were cut after nuisance regression and interpolated using cubic splines.

Independent component analysis (ICA) of resting state fMRI data. ICA was performed using the GIFT toolbox[43]. 20 independent components were estimated using the Infomax algorithm and the ICASSO approach, with a gray substance mask derived from DARTEL segmentation (see above). Networks were labelled based on inspection.

Ancillary outcomes recorded during fMRI scanning

Eye-tracking. Eye-tracking was performed using a Viewpoint eye-tracking system (Arrington Research), mounted in the goggles with which we presented images to the participants. We tracked the right eye only. The objectives of eye-tracking were to verify during scanning that the participants' eyes were open and to investigate pupil width as an indicator of emotional response.

Heart rate. Heart rate was recorded using the pulse oximeter attached to the MRI scanner, for the purpose of investigating it as an indicator of emotional response.

Electromyography (EMG). EMG was recorded using pre-gelled circular 1 cm-diameter radiotranslucent electrodes on 3.8 cm circular vinyl backing. Electrodes were placed over the major zygomatic and superciliary corrugator muscles according to established guidelines[44]. Radiotranslucent clip leads were connected through a patch panel connector to Biopac EMG amplifiers in the control room. Following a recently proposed method[45] to remove scanner noise, EMG signals were processed using a comb band stop filter with a base frequency corresponding to the number of slices/TR. Furthermore, a 30-300 Hz band pass filter was applied to exclude electrical activity not originating in muscle, and a 49-51 Hz band stop filter was used to remove line frequency noise. Average rectified EMG signal was calculated for bins of 1 s during each block. Technical validation of EMG data is not shown in this paper.

Behavioral experiments

Psychomotor vigilance test (PVT). As an indicator of sleepiness, participants completed a 5-minute psychomotor vigilance task before MRI scanning. We used a PVT-192 unit from Ambulatory Monitoring, Inc (Ardsley, NY, USA).

Working memory test. A working memory test was administered to a subset of participants using the WakeApp web app, developed by John Axelsson, on a touchpad (iPad, Apple).

Zürich prosocial game. The Zürich prosocial game[46] was administered to a subset of participants in an attempt to study prosocial behavior.

Blood sampling and biobanking

At the end of each visit, blood was sampled by venipuncture. 5 ml blood for DNA extraction were sampled in one EDTA tube, and 10 ml blood for plasma extraction in another. The 10 ml tube was allowed to rest for 15 minutes at room temperature and then centrifuged for 10 minutes at 3500 rpm and 20°C. The plasma supernatant was aliquoted into 4 cryotubes, which were frozen overnight at -20°C. The 5 ml tube was allowed to stand in room temperature overnight. The next morning, the samples were transferred to the KI Biobank (http://ki.se/en/research/ki-biobank), where plasma samples were frozen at -80°C.

DNA analyses

DNA extraction. DNA was extracted from whole blood samples by KI Biobank using the Chemagic Star DNA Blood 400 kit (Chemagic, Inc.) following the manufacturer's instructions. Briefly, cells were lysed and magnetic beads were used to bind nucleic acids, which were then magnetically separated from other sample material.
Genotyping. DNA analyses were performed on samples from the first session, if available, and from the second session if samples were available from the second session but not the first. Genotyping of the single nucleotide polymorphisms (SNPs) rs4680, rs7410 and rs429358 was performed using the TaqMan SNP genotyping assay on an ABI PRISM 7900 HT Sequence Detection System (Applied Biosystems). Relative telomere length (TL) was determined as previously described[47]. In brief, triplicate DNA samples (4.0 ng) were used both for the telomere (Tel) and the single-copy gene (S) (hemoglobin-b, *HGB*) PCRs performed within the same 384-well plate, amplified by using Power SYBR Green on an ABI PRISM 7900 HT Sequence Detection System. Relative telomere to single copy gene (Tel/S) ratios were determined using a standard curve. In total 2 plates were analyzed, each with three inter-plate calibrator samples included. The primer sequences were (written 5'→3'): Tel1: CGGTTTGTTTGGGTTTGGGTTTGGGTTTGGGTTTGGGTT; Tel2: GGCTTGCCTTACCCTTACCCTTACCCTTACCCTTACCCT; *HGB* Fw: GCTTCTGACACAACTGTGTTCACTAGC; *HGB* Rv: CACCAACTTCATCCACGTTCACC.

Study pre-registration and ethical review
This study was registered to clinicaltrials.gov (NCT02000076). A list of hypotheses was uploaded to the Open Science Framework (https://osf.io/bxfsb/), and subsequently revised as indicated in the submission. The study was approved by the Regional Ethics Review board of Stockholm (2012/1098-31/2), specifically including permission for open publication of de-identified data.

Code availability
Stimulus presentation. Code for stimulus presentation was written in Presentation (Neurobehavioral Systems, Inc.), and is available at https://github.com/GNilsonne/SleepyBrain-StimulusPresentation, under a CC0 licence.
Data analysis. Code for data analyses was written in R and Matlab, and is available at https://github.com/GNilsonne/SleepyBrain-Analyses/tree/master/Resting%20State, including ICA network results, under a CC0 licence.

Data Records
MRI and PSG data
Resting state imaging data, ancillary recordings, genotype and telomere length data, and behavioral data, except from the Zürich Prosocial Game, for which the data will be published at a later time, have been deposited in the openfmri.org database (Data citation 1).

PPI-R item-level responses
Data on the psychopathy personality inventory-revised have been used in a separate project for psychometric validation of this scale in Swedish (Sörman et al, in review). In connection to that publication, item-level data were deposited to Dryad Digital Repository (Data citation 2).

Technical Validation
Participants, randomization, blinding, and efficacy of intervention
Recruitment and inclusion of younger and older participants is shown in figures 1e and 1f, respectively. After participant exclusion due to a finding when the participant arrived to the MRI scanner that inclusion criteria were not fulfilled, and due to incidental MRI findings, the remaining sample was reasonably well counterbalanced with respect to intervention order with 22 vs 27 young and 21 vs 20 old participants starting with the sleep deprivation vs the full

sleep condition. One older participant withdrew from the experiment after session 1. For most outcomes, further attrition occurred due to incomplete data.

We attempted to blind experimenters by the MR scanner to the participants' sleep deprivation status, but this proved a vain hope. Experimenters' guesses of deprivation status were treated as scores of 1-5 and compared between conditions with a 2-sided paired *t*-test, showing a difference of 1.95 [95% CI 1.51, 2.38], $t_{73} = 8.98$, $p < 0.0001$ (figure 1g).

Polysomnography revealed that participants slept shorter in the sleep deprivation condition (figure 1h). We decided on the following criteria for the intervention to have been successful: total sleep time > 4 h in the full sleep condition and < 4 h in the sleep deprivation condition, and a difference between the two conditions of at least 2 h, using manually scored PSG data, and participants' sleep diaries in the cases where PSG data were not available. Nine participants did not meet these criteria. Detailed analyses of polysomnography results will be reported in a forthcoming paper (Åkerstedt et al, in preparation).

Rating scales

There were no missing responses for any forms administered online. All responses made on paper were entered into spreadsheets and double-checked by another investigator. For the sleep diaries, 765 nights were reported by 87 participants. 44 errors were found and corrected out of 32 130 entries (0.1 %). The PPI-R was completed by 85 participants. 15 responses were missing in 10 forms (out of 13 090 total responses, 0.1%). These were judged to be missing completely at random and were imputed from subscale means. One error was found and corrected. For the BIS-11, 40 participants responded online and 40 using the paper version of the scale. Two responses were missing in 2 forms (out of 1 200 total responses on paper, 0.2%), which were judged to be missing completely at random and were imputed from subscale means. No errors were found on double-checking. The BADD scales were completed by 80 participants. Seven responses were missing in 4 forms (out of 3 200 total responses on paper, 0.2%). Of these, 4 were judged to be missing completely at random and were imputed from subscale means. The last 3 were missing from the end of one form and were judged to not be missing completely at random, and were not imputed. One error was found and corrected. PANAS ratings from MRI scanning were missing from two participants, who ended prematurely.

Ancillary outcomes recorded during fMRI scanning

Heart rate. Technical validation was performed on data from the HANDS experiment. Data were recorded from 161 sessions in 86 participants. Heart rate was determined based on recorded pulse events and was investigated within a time window of 4 s before each stimulus onset to 10 s after stimulus onset. Time courses were inspected for each participant and recordings judged as excessively noisy were excluded (n = 22, 14%). Heart rates < 40 beats per minute (bpm) or > 100 bpm were considered non-physiological and were censored. Recordings with less than 50% of data remaining were excluded at this stage (n = 0). Where possible, censored data were imputed using carry-forward of the last non-censored heart rate, which is likely to be a conservative procedure when investigating event-related responses, since it will tend to underestimate changes. 4% of data were censored, out of which 78% were imputed. A mean time course across stimuli and conditions was inspected (figure 1i), showing heart rate acceleration during anticipation, followed by deceleration shortly after stimulus onset. With 4578 events from 115 sessions in 68 participants, a nested mixed-effects model showed that the lower average heart rate during the two seconds following stimulus presentation compared to the index period of -4 to 0 seconds before stimulus onset was not statistically significant (0.997 [95% CI 0.994, 1.001]).

Eye-tracking. Technical validation of pupil diameter measures was performed on data from the HANDS experiment. Data were recorded from 166 sessions in 87 participants. To remove eyeblinks and episodes when the tracker lost the pupil, all records of pupil height and width where the first derivative was < -3 or > 3 were discarded, along with one consecutive data point before and after. Furthermore, all records of pupil height and width < 0.1 cm and > 0.3 cm were discarded. If at least 50 % of data remained in a window from 4 seconds before each event onset to 10 seconds after, a loess curve was fitted to impute the missing data and down-sample the time-course. This yielded X remaining events out of Y. Pupil height and width were averaged to yield a pupil diameter measure. A mean time course across stimuli and conditions was inspected (figure 1j), showing pupil constriction while stimuli were shown. With 2507 events from 102 sessions in 73 participants, a nested mixed-effects model showed that pupil constriction during the 3.5 seconds of stimulus presentation compared to the index period of -4 to 0 seconds before stimulus onset was statistically significant (-0.016 cm [95% CI -0.019, -0.014]).

MRI imaging

Participants' wakefulness during scanning. Two participants presented eye closures > approx. 5 seconds, prompting the scanner operator to give a wake-up call. 26 participants reported at debriefing a subjective experience of nodding off or falling asleep during any of the fMRI experiments.

Quality control by inspection of images. Quality of MRI images was investigated by inspection of anatomical and functional images after conversion from DICOM to NIFTI. The quality control report has been made available with the MRI images. Quality was generally acceptable and no images were excluded based on this screening. One participant had signal dropout affecting the temporal lobe, probably due to some metal object still on the head, and should be excluded specifically from analyses focusing on the amygdala.

Analyses of head movement. Sleep deprivation did not cause considerably more volumes to fall above the threshold for exclusion of framewise displacement > 0.5 mm in the first run (1.2 [95% CI -0.7, 3.0], $p = 0.20$, figure 2a-b) nor in the second run (-0.3, [-2.9, 2.2], $p = 0.79$, figure 2c-d). More volumes exceeded the threshold in the second run compared to the first, both in the full sleep condition (3.8 [2.1, 5.6], $p < 0.0001$) and in the sleep deprivation condition (2.3 [0.4, 4.2], $p = 0.02$). Since 0.5 mm is an arbitrary threshold, we also show in figure 2e-f the mean framewise displacement across all volumes in each run (before interpolation). These analyses suggest that head motion is not a major confounder between sleep deprivation and full sleep conditions.

Resting state networks. Using independent component analysis with 20 components, we successfully identified canonical resting state networks (figure 3).

Genetic analyses

DNA analyses. All DNA samples provided SNP and TL data. All SNPs were in Hardy Weinberg equilibrium ($p > 0.3$). The intra-assay and inter-assay coefficients of variation (CV) for TL measurements were calculated based on the Ct_{Tel} and the Ct_{HGB} values, separately, from these DNA samples detected in two 384-well plates. Intra-assay CV was 0.3% and inter-assay CV was 9%.

Adverse events and incidental findings

Adverse events. Three participants interrupted scanning due to feelings of anxiety or panic. One participant vomited after scanning. In response to an open question at debriefing, three participants described pain during the experiment, one participant described an experience of considerable discomfort, one participant described nausea, one participant reported anxiety during scanning, four participants described the scanner noise as moderately or strongly

unpleasant, and one participant reported unpleasant recollections between the first and second session of emotional stimuli from ARROWS experiment.

Incidental MRI findings. All participants' T2-weighted scans were reviewed by the clinical neuroradiology service at the Karolinska University Hospital. Four participants were excluded because of incidental findings requiring follow-up, which was managed through the Karolinska University Hospital. Additionally, two participants, both in the older group, were recommended to consult their general practitioners for a blood pressure check-up because of a finding of white matter changes. Outcomes of clinical follow-up are not known to us and it is possible that these participants may have had a net benefit or harm, considering the risks involved in additional diagnostic procedures and possible overtreatment. Additionally, 32 participants had findings deemed not to require follow-up, most of which were unspecific white matter changes among the older participants.

Usage Notes

This multimodal dataset can be used to investigate effects of sleep deprivation as well as cross-sectional relationships between various participant characteristics and brain imaging parameters. Samples of serum and DNA are available for additional analyses. Thus, this dataset could be used for discovery or independent validation of associations found in other datasets. Possibly identifying features and variables have been censored in order to safeguard participants' integrity. We encourage researchers to use the published dataset freely and ask that they cite the respective data sources as well as this paper. For any analyses involving censored data, our intention is to share data with colleagues by arrangement, limited in each particular case by regulations pertaining to data integrity as well as by the participants' consent, and we encourage researchers to contact us.

Acknowledgements

We are grateful to Diana Cortes, Danielle Cosme and Roberta Nagai for assistance with polysomnography recordings, to Birgitta Mannerstedt Fogelfors for assistance with screening, instructions to participants, and blood sampling, to Rouslan Sitnikov and Jonathan Berrebi for assistance with MRI sequences and auxiliary equipment, to Hannes Ingre for entering sleep diary data into a spreadsheet, to William Triplett, data curator at openfmri.org, for assistance with data archiving and generating the ISA-Tab metadata file, and to the staff at KI Biobank for assistance with biobanking and DNA extraction.

Author contributions

GN participated in experimental design, recruited and screened participants, collected MRI data, collected blood samples, collected participant ratings and behavioral data, performed curation and quality control, analysed data, deposited data to repositories, and drafted the paper.

ST participated in experimental design, recruited and screened participants, collected MRI data, collected blood samples, collected participant ratings and behavioral data, performed curation and quality control, and analysed data.

PdO collected PSG data, performed curation and quality control of PSG data, and analysed PSG data.

HT collected MRI data, collected participant ratings and behavioral data, performed curation and quality control, and analysed data.

JS participated in experimental design.

CL participated in planning and design of genetic analyses and performed genetic analyses.

JJL performed genetic analyses.

KM participated in planning and design of genetic analyses.

TS participated in experimental design and contributed the WakeApp software.

JA participated in experimental design and contributed the WakeApp software.

PP participated in experimental design and supervised data collection, quality control, and analyses.
PF participated in quality control and analyses of resting state fMRI data.
GK participated in experimental design and supervised data collection, quality control, and analyses.
HF participated in experimental design and supervised data collection, quality control, and analyses.
ML participated in experimental design and supervised data collection, quality control, and analyses.
TÅ participated in experimental design and supervised data collection, quality control, and analyses, and analysed PSG data.

Competing interests

The authors have no competing interests to declare.

Figures

Figure 1. **Participants, randomization, blinding, efficacy of intervention, and heart rate and pupil width responses**

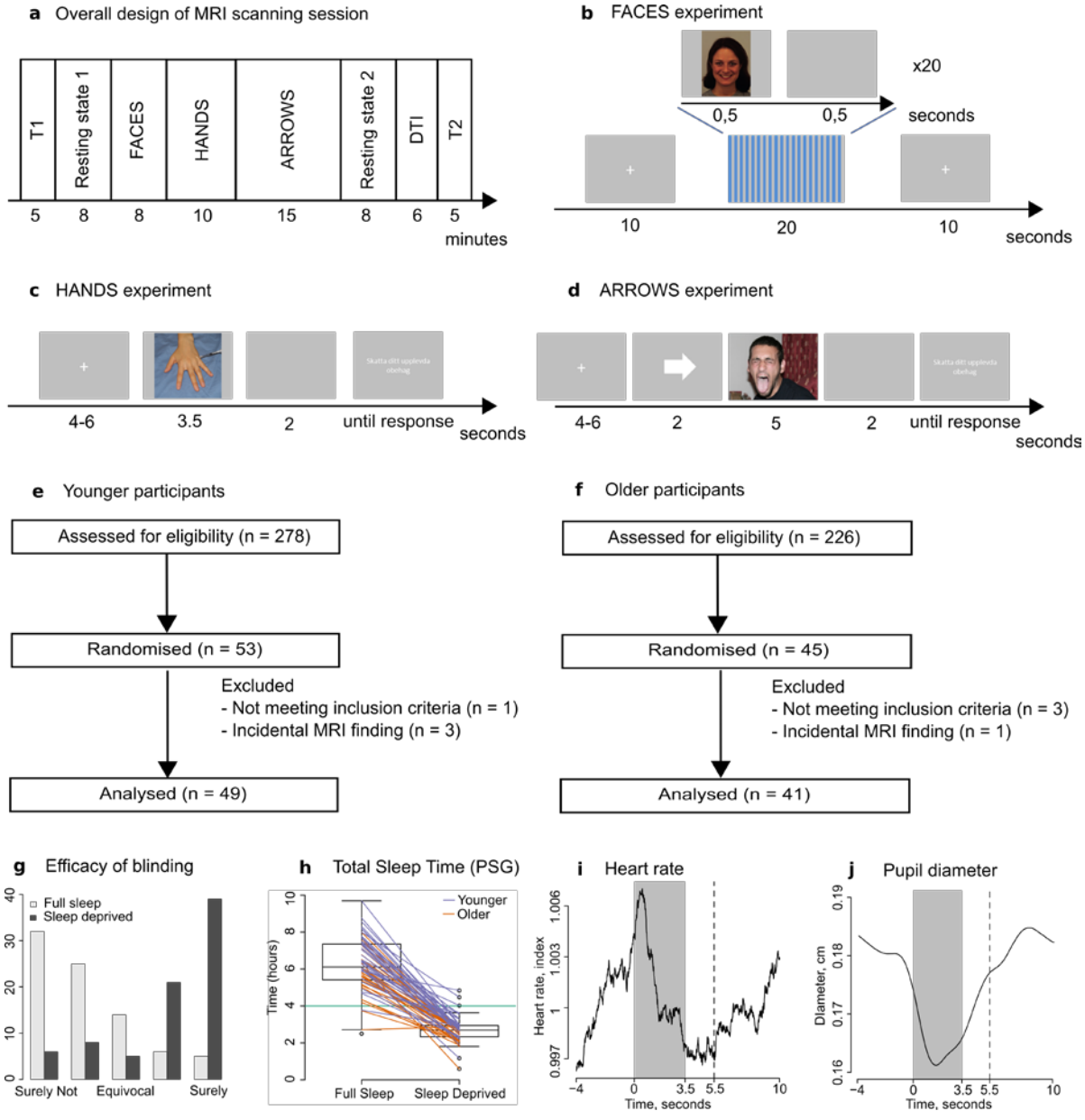

Figure 2. Movement parameters

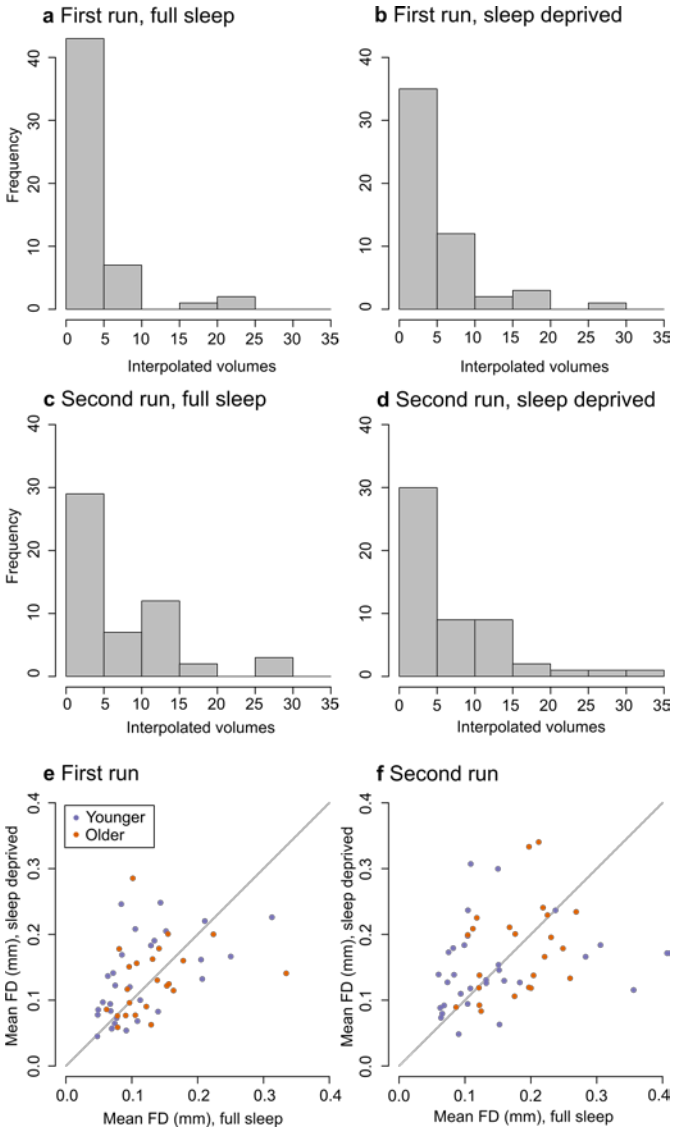

Figure 3. Resting state functional networks

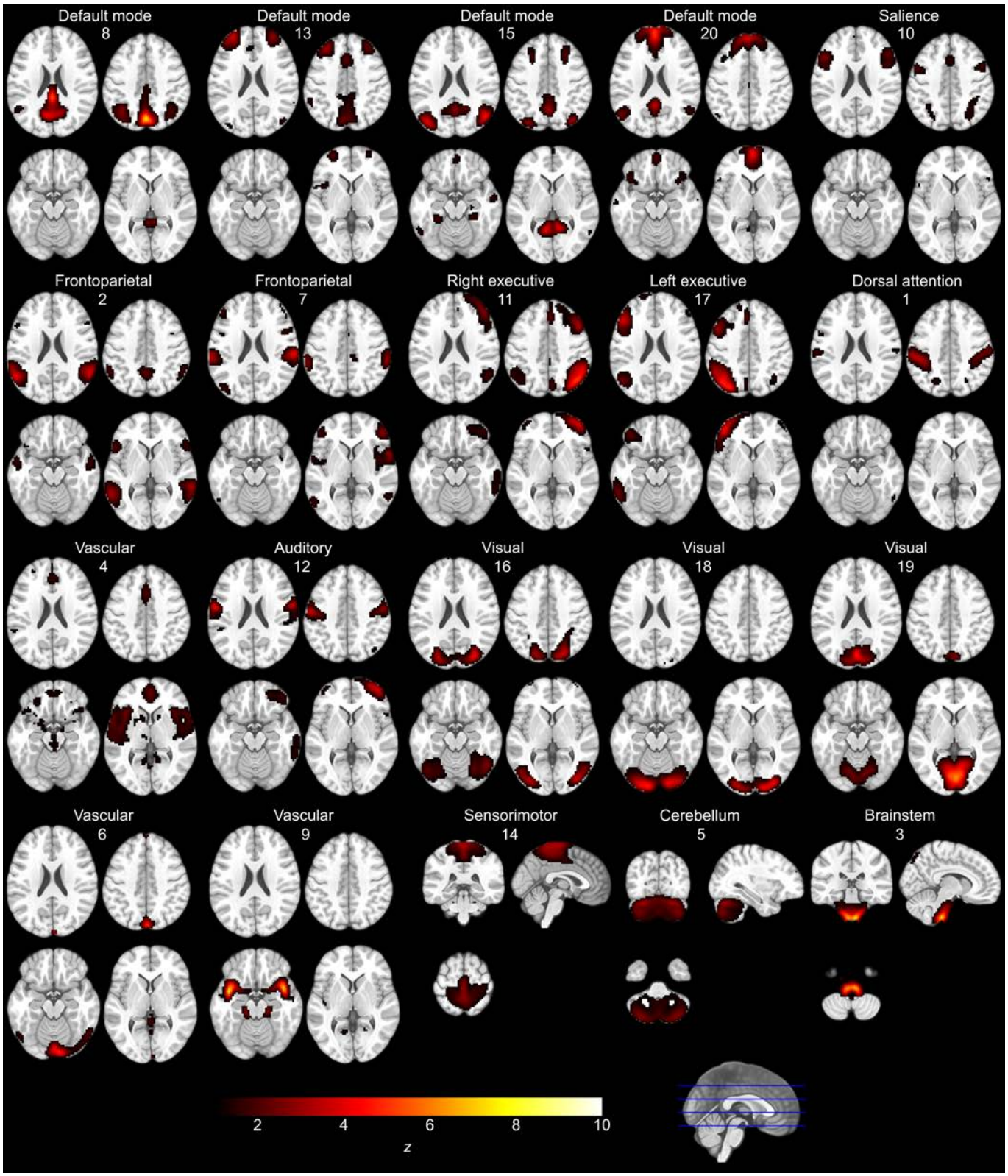

Figure Legends

Figure 1. **Participants, randomization, blinding, efficacy of intervention, and heart rate and pupil width responses**

a: Overall experimental timeline. Karolinska Sleepiness Scale (KSS) ratings were made between each **b:** FACES experiment. **c:** HANDS experiment. **d:** ARROWS experiment. **e:** Inclusion of younger participants. **f:** Inclusion of older participants. **g:** Efficacy of blinding. The experimenter operating the MR scanner rated their suspicion of which condition the participant was in, using a 5-step scale (surely not sleep deprived – likely not sleep deprived – equivocal – likely sleep deprived – surely sleep deprived). For reasons of space, only three of these five labels are shown. **h:** Total sleep time, as measured with polysomnography. The threshold of 4 h for successful intervention is shown with a dashed horizontal line. **i:** Heart rate. Gray box shows time when stimulus was displayed. Vertical dashed line at 5.5 seconds shows onset of rating event. The curve is averaged over all events. **j:** Pupil diameter.

Figure 2. Movement parameters

a-d: Number of volumes interpolated because framewise displacement exceeded 0.5 mm. Participants with more than 40 volumes exceeding this threshold in any run were excluded. **e-f**: Mean framewise displacement across runs.

Figure 3. Resting state functional networks

Of 20 identified networks, 17 were judged to represent blood flow in neural tissues and three were judged to represent vascular components. Networks defined as interesting for comparisons between full sleep and sleep deprivation conditions were the default mode network, the salience network, frontoparietal networks, and executive control networks. These networks were all present and were distributed across nine components. Numbers given are the arbitrary component numbers from independent component analysis. Visualisation was performed with MRIcron with the MNI192 template. Note that the template has higher resolution than the superimposed fMRI results.

Tables

Table 1. ISA-Tab metadata record

See associated Metadata Record.

References

1. Walker, M. P. The Role of Sleep in Cognition and Emotion. *Annals of the New York Academy of Sciences* **1156,** 168–197 (2009).

2. Goldstein, A. N. & Walker, M. P. The role of sleep in emotional brain function. *Annu Rev Clin Psychol* **10,** 679–708 (2014).

3. Gujar, N., Yoo, S.-S., Hu, P. & Walker, M. P. Sleep Deprivation Amplifies Reactivity of Brain Reward Networks, Biasing the Appraisal of Positive Emotional Experiences. *J. Neurosci.* **31,** 4466–4474 (2011).

4. Sämann, P. G. *et al.* Increased sleep pressure reduces resting state functional connectivity. *MAGMA* **23,** 375–389 (2010).

5. Kaida, K. *et al.* Validation of the Karolinska sleepiness scale against performance and EEG variables. *Clinical Neurophysiology* **117,** 1574–1581 (2006).
6. Bastien, C. H., Vallières, A. & Morin, C. M. Validation of the Insomnia Severity Index as an outcome measure for insomnia research. *Sleep Medicine* **2,** 297–307 (2001).
7. Trott, E. *En reliabilitets- och validitetsstudie av den svenska versionen av självskattningsformuläret Insomnia Severity Index (ISI)*. (2009).
8. Zigmond, A. S. & Snaith, R. P. The Hospital Anxiety and Depression Scale. *Acta Psychiatrica Scandinavica* **67,** 361–370 (1983).
9. Lisspers, J., Nygren, A. & Söderman, E. Hospital Anxiety and Depression Scale (HAD): some psychometric data for a Swedish sample. *Acta Psychiatrica Scandinavica* **96,** 281–286 (1997).
10. Kecklund, G. & Åkerstedt, T. The psychometric properties of the Karolinska Sleep Questionnaire. *J Sleep Res* **1,** 113 (1992).
11. Davis, M. A multidimensional approach to individual differences in empathy. *JSAS Catalog of Selected Documents in Psychology* **10,** (1980).
12. Davis, M. Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of Personality and Social Psychology* **44,** 113–126 (1983).
13. Cliffordson, C. The hierarchical structure of empathy: Dimensional organization and relations to social functioning. *Scandinavian Journal of Psychology* **43,** 49–59 (2002).
14. Bagby, R. M., Parker, J. D. A. & Taylor, G. J. The twenty-item Toronto Alexithymia scale—I. Item selection and cross-validation of the factor structure. *Journal of Psychosomatic Research* **38,** 23–32 (1994).
15. Simonsson-Sarnecki, M. *et al.* A Swedish Translation of the 20-item Toronto Alexithymia Scale: Cross-validation of the Factor Structure. *Scandinavian Journal of Psychology* **41,** 25–30 (2000).
16. Spielberger, C. D. in *Corsini Encyclopedia of Psychology* (John Wiley & Sons, Inc., 2005).

17. Benet-Martínez, V. & John, O. P. Los Cinco Grandes across cultures and ethnic groups: Multitrait-multimethod analyses of the Big Five in Spanish and English. *Journal of Personality and Social Psychology* **75,** 729–750 (1998).
18. John, O. P., Donahue, E. M. & Kentle, R. L. The Big Five Inventory--Versions 4a and 54. (1991).
19. John, O. P., Naumann, L. P. & Soto, C. J. in *Handbook of personality: Theory and research (3rd ed.).* 114–158 (Guilford Press, 2008).
20. Zakrisson, I. *Big Five Inventory (BFI) [Elektronisk resurs] : Utprövning för svenska förhållanden*. (Mittuniversitetet, 2010).
21. Cohen, S., Kamarck, T. & Mermelstein, R. A Global Measure of Perceived Stress. *Journal of Health and Social Behavior* **24,** 385 (1983).
22. Torsvall, L. & Akerstedt, T. A diurnal type scale. Construction, consistency and validation in shift work. *Scand J Work Environ Health* **6,** 283–290 (1980).
23. Axelsson, J., Akerstedt, T., Kecklund, G. & Lowden, A. Tolerance to shift work-how does it relate to sleep and wakefulness? *Int Arch Occup Environ Health* **77,** 121–129 (2004).
24. Johns, M. W. A new method for measuring daytime sleepiness: the Epworth sleepiness scale. *Sleep* **14,** 540–545 (1991).
25. Watson, D., Clark, L. A. & Tellegen, A. Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology* **54,** 1063–1070 (1988).
26. Hillerås, P. K., Jorm, A. F., Herlitz, A. & Winblad, B. Negative and Positive Affect among the Very Old A Survey on a Sample Age 90 Years or Older. *Research on Aging* **20,** 593–610 (1998).
27. Eriksson, I., Undén, A.-L. & Elofsson, S. Self-Rated Health. Comparisons Between Three Different Measures. Results from a Population Study. *Int. J. Epidemiol.* **30,** 326–333 (2001).

28. Patton, J. H., Stanford, M. S. & Barratt, E. S. Factor structure of the Barratt impulsiveness scale. *J Clin Psychol* **51,** 768–774 (1995).
29. Åkerstedt, T., Hume, K., Minors, D. & Waterhouse, J. Good sleep — its timing and physiological sleep characteristics. *Journal of Sleep Research* **6,** 221–229 (1997).
30. Lilienfeld, S. O. & Andrews, B. P. Development and Preliminary Validation of a Self-Report Measure of Psychopathic Personality Traits in Noncriminal Population. *Journal of Personality Assessment* **66,** 488–524 (1996).
31. TE, B. Brown attention-deficit disorder scales. *San Antonio, TX: The Psychological Corporation* (1996).
32. Brown, T. E. & Järvå, H. *Brown attention-deficit disorder scales® for children and adolescents: svenskt manualsupplement*. (Psykologiförl., 2004).
33. Folstein, M. F., Folstein, S. E. & McHugh, P. R. 'Mini-mental state': A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research* **12,** 189–198 (1975).
34. Folstein, M. F., Folstein, S. E. & Fanjiang, G. *MMSE Mini-Mental State Examination. Clinical Guide*. (Psychological Assessment Resources, Inc., 2001).
35. *AASM manual for the scoring of sleep and associated events*. (American Academy of Sleep Medicine, 2007).
36. Lundqvist, D., Flykt, A. & Öhman, A. *The Karolinska Directed Emotional Faces-KDEF. CD-ROM from Department of Clinical Neuroscience, Psychology section, Karolinska Institutet, Stockholm, Sweden*. (ISBN 91-630-7164-9, 1998).
37. Lamm, C., Nusbaum, H. C., Meltzoff, A. N. & Decety, J. What Are You Feeling? Using Functional Magnetic Resonance Imaging to Assess the Modulation of Sensory and Affective Responses during Empathy for Pain. *PLoS ONE* **2,** e1292 (2007).
38. Lang, P. J., Bradley, M. M. & Cuthbert, B. N. International affective picture system (IAPS): Affective ratings of pictures and instruction manual. *Technical report A-8* (2008).

39. Golkar, A. *et al.* Distinct Contributions of the Dorsolateral Prefrontal and Orbitofrontal Cortex during Emotion Regulation. *PLoS ONE* **7,** e48107 (2012).

40. Chao-Gan, Y. & Yu-Feng, Z. DPARSF: A MATLAB Toolbox for 'Pipeline' Data Analysis of Resting-State fMRI. *Front Syst Neurosci* **4,** (2010).

41. Ashburner, J. A fast diffeomorphic image registration algorithm. *Neuroimage* **38,** 95–113 (2007).

42. Power, J. D., Barnes, K. A., Snyder, A. Z., Schlaggar, B. L. & Petersen, S. E. Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *Neuroimage* **59,** 2142–2154 (2012).

43. Calhoun, V. D., Adali, T., Pearlson, G. D. & Pekar, J. J. A method for making group inferences from functional MRI data using independent component analysis. *Hum Brain Mapp* **14,** 140–151 (2001).

44. Fridlund, A. J. & Cacioppo, J. T. Guidelines for Human Electromyographic Research. *Psychophysiology* **23,** 567–589 (1986).

45. Heller, A. S., Greischar, L. L., Honor, A., Anderle, M. J. & Davidson, R. J. Simultaneous acquisition of corrugator electromyography and functional magnetic resonance imaging: A new method for objectively measuring affect and neural activity concurrently. *NeuroImage* **58,** 930–934 (2011).

46. Leiberg, S., Klimecki, O. & Singer, T. Short-Term Compassion Training Increases Prosocial Behavior in a Newly Developed Prosocial Game. *PLoS ONE* **6,** e17798 (2011).

47. Wei, Y. B. *et al.* hTERT genetic variation in depression. *J Affect Disord* **189,** 62–69 (2016).

Data Citations

Bibliographic information for the data records described in the manuscript.

1. Nilsonne, G. & al. openfrmi.org ds000201 (https://openfmri.org/dataset/ds000201/).
2. Sörman, K., Nilsonne, G., Howner, K., Tamm, S., Caman, S., Wang, H., Edens, J.F., Gustavsson, P., Lilienfeld, S.O., Petrovic, P., Fischer, H. & Kristiansson, M. Data from: Reliability and construct validity of the Psychopathic Personality Inventory-Revised in a Swedish non-criminal sample. Dryad Digital Repository. http://dx.doi.org/10.5061/dryad.qh8c9 (2016)