**SUPPLEMENTAL MATERIALS**

**Supplemental Methods**

*Circadian phase*

Circadian phase was assessed via salivary Dim Light Melatonin Onset (DLMO) (Benloucif et al., 2008) during the second evening of each Aligned and Misaligned study visit. Saliva samples were collected in Salivettes (Sarstedt, Newton, NC) under dim light conditions (<15 lux at any angle of gaze) every half-hour from 1800 - 0100. Dim light conditions began at 1730 and were confirmed using a light meter. The sampling protocol controlled for posture (participants remained seated other than trips to the bathroom) and other confounding factors (e.g., no eating/drinking within 10 min of sampling; caffeine, bananas, and chocolate prohibited) (Burgess, 2010). Participants rinsed their mouths with water 10 minutes prior to each sample after eating or drinking. For DLMO collection during the Aligned condition, the fMRI scan (1900-1930) occurred after saliva sample collection began (1800). Therefore, participants donned welder’s goggles (Uvex Flex Seal with Shade 5.0 Infra Dura lens; illuminance <0.25 lux) (Burgess, 2010) during transportation to and from the scan center and wore fMRI-compatible blue blocking goggles (Traditionalist model by LowBlueLights from cet.org) during the scan. Saliva samples were frozen at -80° C and shipped on dry ice for radioimmunoassay (Solid Phase, Inc.; Portland, ME) using commercially available kits (ALPCO, Inc., Salem, NH). DLMO was calculated as the interpolated clock time when salivary melatonin levels (pg/mL) exceeded the mean of three consecutive baseline samples plus twice the standard deviation of those samples (Voultsios, Kennaway, & Dawson, 1997).

*fMRI Tasks*

As noted, participants completed a total of four fMRI scans. Given concerns over habituation to similar reward tasks during test-retest intervals of less than 2 weeks, we only administered the Card Guessing Task during the AM scans (Hasler, Forbes, & Franzen, 2014; Plichta et al., 2012).

Taskswere presented by projecting images onto a rear projection screen at the subject’s chest level and viewed through a mirror attached to the head coil. Stimulus presentation and registration of responses were controlled by a Windows-based computer running E-Prime 2.0 (Psychology Software Tools, 2012). Participants’ button press responses and reaction times were recorded using an RF-shielded response box and cable connected to the computer.

*fMRI Card Guessing Task*. We employed a slow event-related fMRI card-guessing paradigm designed to examine neural reactivity to anticipation and receipt of monetary reward and loss (see (Hasler et al., 2013)). Trials were presented in a pseudorandom order with predetermined outcomes. Each 20-s trial consisted of a 4-s decision phase when participants guessed whether the value of a visually presented card with a possible value of 1–9 would be higher or lower than 5; a 6-s anticipation period when the trial type (possible-reward or possible-loss) was displayed; a 1–s outcome period the numerical value of the card and then outcome feedback (win, loss, or no-change) were presented; and a 9-s interstimulus interval. In reward trials, participants were told they would win $5 if their guess was correct and there would be no change in earnings if their guess was incorrect. In loss trials, participants were told they would lose $1 if their guess was incorrect and there would be no change in earnings if their guess was correct. Twenty-four trials were presented in one run with 12 reward-anticipation and 12 loss-anticipation trials. Within reward-anticipation trials there were a balanced number of win-outcome and no-change outcome trials. Outcome probabilities were fixed trial-wise to ensure an identical win/loss time series modeling and pattern of outcome experiences for every participant. Participants were unaware of the fixed outcome probabilities in the paradigm and were led to believe their performance would determine net monetary gain. Each participant was given $25 in earnings.

In the present study we focus on the anticipation and outcome phases of the reward trials, based on evidence that circadian preference is more strongly related to reward, appetitive motivation, and positive affect than to negative affect processes (Hasler, Allen, Sbarra, Bootzin, & Bernert, 2010). Consistent with our prior work using this task (e.g., (Hasler et al., 2012; Hasler et al., 2013), we focused on the Reward Anticipation>Baseline and Reward Win>Baseline contrasts, in which baseline was defined as the last 3 s of the interstimulus interval.

*fMRI Go/No-Go Task*. This variant of a simple and widely-used Go/No-Go paradigm (Casey et al., 1997) was used to probe response inhibition. Participants are shown a continuous series of 120 letters and instructed to respond to any letter except V (V was used as the non-target instead of X because X may be associated with implicit meanings such as “Stop!”). Seventy-five percent of the trials are targets (i.e. letters other than V). The task included two conditions: Block A—the Go condition has 20 targets, and Block B—the NoGo condition has 10 targets and 10 nontargets. Within each 30-second block, targets are presented in pseudo-randomized order. Blocks are presented in the order ABBABA, rather than ABABAB in order to reduce the predictability of the task. A 20-second rest block followed each round of Blocks A and B. Stimulus presentation duration is 0.5 seconds and inter-stimulus interval was 1.5 seconds, resulting in a total of 120 stimuli presented in a 290 s period (including rest blocks). Participants were not given an indication of their performance during the task. The contrast of interest was No-Go > Rest in order to capture the neural correlates of active response inhibition. Behavioral outcomes included reaction time during the Go condition and accuracy during the No-Go condition.

*Neuroimaging acquisition*

Neuroimaging was conducted on a Siemens 3T TIM Trio scanner (years 1-3 of the grant) and a Siemens 3T Biograph mMR scanner (years 4 and 5 of the grant). Functional images were acquired using a gradient echo planar imaging (EPI) sequence that included 39 axial slices (3.1 mm wide) beginning at the cerebral vertex and extending across the entire cerebrum and most of the cerebellum (TR/TE = 2000/25 ms, field of view = 205 mm, matrix = 64 × 64). Scanning parameters were selected to optimize BOLD signal quality while maximizing whole brain coverage. A reference EPI scan was acquired before fMRI data collection to visually inspect for artifacts (e.g., ghosting) and ensure adequate signal across the entire volume. In addition, a 192-slice high-resolution axially-acquired T1-weighted anatomical image was collected for co-registration and normalization of functional images.

**Neuroimaging preprocessing and analysis**

Neuroimaging data were preprocessed and analyzed using SPM12 (Ashburner et al., 2014). Preprocessing steps included realignment each participant’s data to the first volume in the time series to correct for head motion, coregistration of the realigned image with the subject’s anatomical image, segmentation of the structural scan to provide the necessary transformation parameters for spatial normalization, normalization of the functional scan to the standard Montreal Neurological Institute (MNI) template using the deformation field calculated during segmentation, and spatial smoothing with a Gaussian kernel of 6 mm full-width at half-maximum. Artifact Detection Toolbox (ART; http://www.nitrc.org/projects/artifact\_detect/) software was used to detect functional volumes with signal intensity >4 SD from the subject's mean or >2mm composite interscan motion. Preprocessed data were inspected prior to second-level analysis to ensure that all scans had fewer than 20% of volumes with excessive movement detected by ART and good scan quality. Temporal censoring based on ART output was used to remove motion artifacts in first-level analysis for the remaining scans (Siegel et al., 2014). Five (out of 49) reward scans and 13 (out of 100) Go/No-Go scans were excluded on this basis.

Preprocessed data were analyzed using a two-level random-effects procedure within a general linear modeling (GLM) framework. For each participant and scan (first-level), the main effect of task at each voxel was calculated for the relevant contrast of interest (e.g., Reward Anticipation>Baseline; No-Go>Rest). Temporal censoring based on ART and the six movement parameters were included as regressors in order to control for participant movement. These first-level contrast images were then included in three second-level 1-sample t-tests of each relevant contrast of interest: Reward Anticipation>Baseline and Win Outcome>Baseline for the Card Guessing Task; No-Go>Rest for Go/No-Go. We used the *rex* toolbox (<https://web.mit.edu/swg/software.htm>) to extract mean activations (p=1.0) across all voxels from *a priori* regions of interest (ROIs). For the Card Guessing Task, we employed bilateral VS and the mPFC ROIs previously demonstrated to be responsive to the task. The VS ROI was constructed using WFU PickAtlas Tool v2.4 and comprised two spheres of 10mm radius centered on MNI coordinates x=+/−12, y=12, z=-10 (L. Murray, Shaw, Forbes, & Hyde, 2017). The mPFC ROI was also constructed using the PickAtlas and defined as a 25-mm radius sphere centered on MNI coordinates x=0, y=44, z=22 and including medial Brodmann Areas (BA) 10 and BA32 (E. E. Forbes et al., 2010). For the Go/No-Go Task, we focused on the right inferior frontal gyrus (IFG) as the region most consistently associated with response inhibition (Bari & Robbins, 2013). We based the rIFG ROI on results a meta-analysis of Go/No-Go studies (Criaud & Boulinguez, 2013), constructing a sphere of 10-mm radius around MNI coordinates x=35, y=21, z=−2. All ROIs are shown in Figure 1. As an additional quality control beyond ART (see above), we excluded scans from analysis with <70% coverage of the ROI (5 bilateral VS and 6 mPFC ROIs among the reward scans); 4 rIFG ROIs among the Go/No-Go scans) and/or with z-scores >3 or <-3, resulting in n=43-44 included reward scans (Aligned and Misaligned combined; all AM) and n=82 included Go/No-Go scans (Aligned and Misaligned, AM and PM combined). Finally, we converted all extracted mean activations to z-scores in order to enhance interpretability.

Figure S1. The three region-of-interest (ROI) masks used in the study, including the bilateral ventral striatum (VS; A), the medial prefontal cortex (mPFC; B), and the right inferior frontal gyrus (rIFG; C).

A picture containing cake, indoor, decorated, food

Description automatically generatedA close up of a flower

Description automatically generated A picture containing cake, food, fruit, birthday

Description automatically generated

**A**

**C**

**B**

**Supplemental Results**

Table S1. Differences in mean actigraphy-based sleep parameters across each of the Stabilization weeks (at home) by condition (Aligned vs Misaligned)

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Aligned | | Misaligned | |  |  |
|  | *M* | *SD* | *M* | *SD* | *t* | *p* |
| Lights out | 00:27 | 00:32 | 00:19 | 00:54 | 0.95 | 0.355 |
| Sleep onset latency | 30 m | 25 m | 30 m | 27 m | -0.16 | 0.872 |
| Sleep onset | 00:57 | 00:38 | 00:49 | 00:52 | 0.74 | 0.470 |
| Wake after sleep onset | 50 m | 17 m | 48 m | 17 m | 0.87 | 0.395 |
| Sleep offset | 8:55 | 00:43 | 8:40 | 00:54 | 1.28 | 0.214 |
| Out of bed | 9:15 | 00:35 | 9:02 | 00:51 | 1.14 | 0.268 |
| Total sleep time | 7 h 8 m | 42 m | 7h 2m | 44 m | 0.85 | 0.405 |
| Time in bed | 8h 48m | 38 m | 8h 43 m | 48 m | 0.69 | 0.495 |
| Sleep efficiency | 81.21 | 7.07 | 80.91 | 7.07 | 0.40 | 0.696 |

Note: based on n=22 participants with both Aligned and Misaligned actigraphy data at home during the 7 nights of stabilization

Table S2. Differences in actigraphy-based sleep parameters during manipulation (laboratory) nights by condition

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Aligned | | Misaligned | |  |  |
|  | *M* | *SD* | *M* | *SD* | *t* | *p* |
| Lights out | 00:11 | 00:27 | 20:21 | 00:41 | 41.09 | **0.000** |
| Sleep onset latency | 23 m | 17 m | 14 m | 13 m | 2.01 | 0.058 |
| Sleep onset | 00:33 | 00:33 | 20:35 | 00:38 | 38.71 | **0.000** |
| Wake after sleep onset | 41 m | 17 m | 69 m | 51 m | -2.62 | **0.017** |
| Sleep offset | 9:06 | 00:23 | 5:11 | 00:28 | 34.08 | **0.000** |
| Out of bed | 9:21 | 00:20 | 5:30 | 00:09 | 60.16 | **0.000** |
| Total sleep time | 7h 51m | 41m | 7h 27m | 1h 0m | 2.88 | **0.010** |
| Time in bed | 9 h 11 m | 31 m | 9 h 8 m | 43 m | 0.37 | 0.718 |
| Sleep efficiency | 85.62 | 6.25 | 81.39 | 8.38 | 2.79 | **0.012** |

Note: based on n=20 participants with both Aligned and Misaligned actigraphy data on lab nights

Table S3. Models predicting bilateral VS activation during reward anticipation versus baseline (standardized score)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Parameter | B | SE | t-value | p-value |
| **Core model** |  |  |  |  |
| Intercept | 3.63 | 3.15 | 1.15 | 0.257 |
| Gender | 0.37 | 0.28 | 1.34 | 0.196 |
| Scanner | 0.49 | 0.29 | 1.67 | 0.110 |
| Scan Visit | -0.21 | 0.20 | -1.06 | 0.304 |
| Scan Order | 0.16 | 0.29 | 0.56 | 0.585 |
| DLMO | -0.20 | 0.15 | -1.32 | 0.195 |
| Align vs Misalign | 0.69 | 0.29 | 2.36 | **0.026** |
|  |  |  |  |  |
| **Core model + lab night TST** | |  |  |  |
| Lab night TST | -0.02 | 0.18 | -0.11 | 0.911 |
| Align vs Misalign | 0.79 | 0.31 | 2.55 | **0.018** |
|  |  |  |  |  |
| **Core model + PVT** |  |  |  |  |
| PVT | 0.04 | 0.06 | 0.66 | 0.513 |
| Align vs Misalign | 0.86 | 0.35 | 2.51 | **0.021** |

NOTE: The models adding lab night TST or PVT include all the covariates included in the Core model, but they are not listed for clarity of presentation. All covariate effects remained non-significant in subsequent models.

Table S4. Models predicting mPFC activation during reward anticipation versus baseline (standardized score)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Parameter | B | SE | t-value | p-value |
| **Core model** |  |  |  |  |
| Intercept | 2.29 | 2.80 | 0.82 | 0.419 |
| Gender | -0.08 | 0.27 | -0.30 | 0.770 |
| Scanner | 0.55 | 0.28 | 1.97 | 0.062 |
| Scan Visit | -0.21 | 0.17 | -1.25 | 0.225 |
| Scan Order | -0.02 | 0.27 | -0.07 | 0.947 |
| DLMO | -0.12 | 0.13 | -0.88 | 0.386 |
| Align vs Misalign | 0.43 | 0.24 | 1.75 | 0.093 |
|  |  |  |  |  |
| **Core model + lab night TST** | |  |  |  |
| Lab night TST | 0.08 | 0.17 | 0.47 | 0.640 |
| Align vs Misalign | 0.40 | 0.30 | 1.35 | 0.192 |
|  |  |  |  |  |
| **Core model + PVT** |  |  |  |  |
| PVT | -0.09 | 0.05 | -1.71 | 0.101 |
| Align vs Misalign | 0.17 | 0.26 | 0.66 | 0.519 |

NOTE: The models adding lab night TST or PVT include all the covariates included in the Core model, but they are not listed for clarity of presentation. All covariate effects remained non-significant in subsequent models.

Table S5. Models predicting bilateral VS activation during reward win versus baseline (standardized score)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Parameter | B | SE | t-value | p-value |
| **Core model** |  |  |  |  |
| Intercept | 2.66 | 2.97 | 0.89 | 0.378 |
| Gender | 0.50 | 0.23 | 2.13 | 0.050 |
| Scanner | 0.23 | 0.25 | 0.91 | 0.377 |
| Scan Visit | 0.08 | 0.22 | 0.37 | 0.718 |
| Scan Order | 0.02 | 0.24 | 0.07 | 0.942 |
| DLMO | -0.15 | 0.14 | -1.03 | 0.310 |
| Align vs Misalign | 0.54 | 0.30 | 1.79 | 0.086 |
|  |  |  |  |  |
| **Core model + lab night TST** | |  |  |  |
| Lab night TST | -0.23 | 0.15 | -1.49 | 0.147 |
| Align vs Misalign | 0.83 | 0.28 | 2.98 | **0.008** |
|  |  |  |  |  |
| **Core model + PVT** |  |  |  |  |
| PVT | 0.08 | 0.05 | 1.53 | 0.140 |
| Align vs Misalign | 0.73 | 0.35 | 2.11 | **0.048** |

NOTE: The models adding lab night TST or PVT include all the covariates included in the Core model, but they are not listed for clarity of presentation. Nearly all covariate effects remained non-significant in subsequent models, with the exception of gender, which became statistically-significant (p<0.05) after accounting for TST or PVT.

Table S6. Models predicting mPFC activation during reward win versus baseline (standardized score)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Parameter | B | SE | t-value | p-value |
| **Core model** |  |  |  |  |
| Intercept | 2.52 | 2.52 | 1.00 | 0.323 |
| Gender | 0.20 | 0.23 | 0.87 | 0.397 |
| Scanner | 0.47 | 0.25 | 1.92 | 0.071 |
| Scan Visit | 0.00 | 0.15 | 0.01 | 0.992 |
| Scan Order | -0.11 | 0.24 | -0.47 | 0.647 |
| DLMO | -0.13 | 0.12 | -1.07 | 0.290 |
| Align vs Misalign | 0.29 | 0.22 | 1.30 | 0.207 |
|  |  |  |  |  |
| **Core model + lab night TST** | |  |  |  |
| Lab night TST | 0.00 | 0.14 | 0.01 | 0.996 |
| Align vs Misalign | 0.37 | 0.25 | 1.50 | 0.149 |
|  |  |  |  |  |
| **Core model + PVT** |  |  |  |  |
| PVT | -0.08 | 0.03 | -2.90 | **0.014** |
| Align vs Misalign | 0.16 | 0.14 | 1.15 | 0.278 |

NOTE: The models adding lab night TST or PVT include all the covariates included in the Core model, but they are not listed for clarity of presentation. All covariate effects remained non-significant in subsequent models.

Table S7. Models predicting rIFG activation during No Go versus Rest (standardized score), comparing all four time points (both AM and PM scans in Aligned and Misaligned conditions)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Parameter** | **B** | **SE** | **t-value** | **p-value** |
| **Core model** |  |  |  |  |
| Intercept | 3.46 | 2.00 | 1.73 | 0.089 |
| Gender | 0.23 | 0.23 | 1.02 | 0.319 |
| Scanner | 0.92 | 0.24 | 3.77 | **0.001** |
| Scan Visit | 0.32 | 0.14 | 2.38 | **0.021** |
| Scan Order | -0.06 | 0.23 | -0.24 | 0.815 |
| DLMO | -0.20 | 0.09 | -2.09 | **0.041** |
| Align vs Misalign | -0.04 | 0.21 | -0.19 | 0.848 |
| AM vs PM | -0.24 | 0.19 | -1.27 | 0.210 |
| Align vs Misalign X AM vs PM | 0.64 | 0.27 | 2.42 | **0.019** |

NOTE: The models adding lab night TST or PVT include all the covariates included in the Core model, but they are not listed for clarity of presentation. All covariate effects remained non-significant in subsequent models.

Table S8. Models predicting rIFG activation during No Go versus Rest (standardized score), comparing AM scans only in Aligned and Misaligned conditions

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Parameter** | **B** | | **SE** | | **t-value** | | **p-value** | |
| **Core model** |  | |  | |  | |  | |
| Intercept | 2.51 | | 2.72 | | 0.92 | | 0.364 | |
| Gender | 0.45 | | 0.27 | | 1.71 | | 0.107 | |
| Scanner | 0.96 | | 0.28 | | 3.40 | | **0.004** | |
| Scan Visit | 0.33 | | 0.19 | | 1.67 | | 0.116 | |
| Scan Order | 0.24 | | 0.27 | | 0.87 | | 0.398 | |
| DLMO | -0.18 | | 0.13 | | -1.38 | | 0.179 | |
| Align vs Misalign | 0.67 | | 0.25 | | 2.69 | | **0.013** | |
|  |  | |  | |  | |  | |
| **Core model + lab night TST** | |  | |  | |  | |
| Lab night TST | -0.17 | | 0.17 | | -1.02 | | 0.320 | |
| Align vs Misalign | 0.86 | | 0.31 | | 2.79 | | **0.012** | |
|  |  | |  | |  | |  | |
| **Core model + PVT** | |  | |  | |  | |
| PVT | 0.01 | | 0.07 | | 0.13 | | 0.899 | |
| Align vs Misalign | 0.87 | | 0.31 | | 2.77 | | **0.012** | |

NOTE: The models adding lab night TST or PVT include all the covariates included in the Core model, but they are not listed for clarity of presentation. All covariate effects remained non-significant in subsequent models.

POWER ANALYSIS

We estimated power for testing the sleep condition (Aligned vs. Misaligned; a 2-level within-subject factor) within a mixed effects model with either 2 repeated measures (fMRI measures of reward) or 4 repeated measures (response inhibition). To reflect the study design, we included another 2-factor within-subject factor to reflect study visit (one versus two; d=0.20) and a 2-factor between subject factor to reflect order (Align-Misalign versus Misalign-Align; d=0.20). We considered type-1 errors of α=0.0125 (adjusted for multiple comparisons) and α=0.05 (no multiple comparison adjustment), assumed N=22 (12 per between-subject group for order), and included a random intercept. Power estimates were computed based on 500 simulations within PASS version 13.0.8. Results are summarized in Table S9.

**Table S9**. Power (1-β) for detecting various effect sizes of the sleep condition (Aligned vs. Misaligned) with 2 to 4 repeated measures. (Based on simulations, some power estimates did not differ with 2 versus 4 time points.)

|  |  |  |
| --- | --- | --- |
| **Effect Size (d)** | **α = 0.05** | **α = 0.0125** |
| 0.40 | 0.47 - 0.48 | 0.28 – 0.28 |
| 0.45 | 0.55 – 0.61 | 0.37 – 0.37 |
| 0.50 | 0.69 - 0.70 | 0.44 - 0.46 |
| 0.55 | 0.74 - 0.78 | 0.53 - 0.56 |
| 0.60 | 0.83 – 0.84 | 0.65 - 0.66 |
| 0.65 | 0.88 - 0.90 | 0.72 - 0.73 |
| 0.70 | 0.92 - 0.93 | 0.81 - 0.82 |
| 0.75 | 0.95 – 0.96 | 0.87 - 0.88 |
| 0.80 | 0.97 - 0.98 | 0.91 – 0.93 |

Prior to adjusting for multiple comparisons (α = 0.05) we estimate that we have at least 0.80 power to detect moderate effects of size d=0.60 and 0.90 power to detect moderate effects of size d=0.70. After adjusting for multiple comparisons (α = 0.0125) we estimate that we have at least 0.80 power to detect moderate effects of size d=0.70 and 0.90 power to detect large effects of size d=0.80. Thus, although the total number of participants is relatively small, by leveraging a within-subjects design where each participant receives both the aligned and misaligned condition, we are sufficiently powered to detect moderate effect sizes prior to multiple comparison adjustment, and moderate-to-large effect sizes after multiple comparison adjustment.

The actual sleep condition effects that we detected with statistical significance (prior to multiple comparison adjustment) ranged from 0.67 to 0.87. As shown by our power analysis, we were sufficiently powered to detect these effects.