Does fMRI neurofeedback in the context of stress influence mood and arousal? A randomised controlled trial with parallel group design [version 1; peer review: awaiting peer review]

Angelo Belardi1, Jong-Hwan Lee2, Hyun-Chul Kim2, Esther Stalujanis1,3, Eun Kyung Jung2, Minkyung Oh2, Seung-Schik Yoo4, Jens C. Pruessner5, Marion Tegethoff1*, Gunther Meinschmidt3,6,7*

1Division of Clinical Psychology and Psychiatry, Department of Psychology, University of Basel, Basel, Switzerland
2Department of Brain and Cognitive Engineering, Korea University, Seoul, South Korea
3Division of Clinical Psychology and Cognitive Behavioral Therapy, International Psychoanalytic University, Berlin, Germany
4Department of Radiology, Brigham and Women’s Hospital, Harvard Medical School, Boston, USA
5Department of Psychology, University of Konstanz, Konstanz, Germany
6Division of Clinical Psychology and Epidemiology, Department of Psychology, University of Basel, Basel, Switzerland
7Department of Psychosomatic Medicine, University Hospital Basel and University of Basel, Basel, Switzerland

* Equal contributors

Abstract

Background: Stress-related mental and physical health issues burden modern societies. New treatment opportunities could help to lessen long-term detrimental consequences of stress.

Objective: To investigate whether real-time functional magnetic resonance imaging neurofeedback (rtfMRInf), aimed at modulating brain activity associated with a stressor, affects subjective mood and arousal.

Methods: In total, 30 males participated in a randomised controlled trial with parallel-group design. rtfMRInf was the intervention, sham-neurofeedback the control condition, and the Stroop task the stressor. We instructed participants to modulate their stress response to the Stroop task via feedback from their anterior cingulate cortex and their insular cortex, concomitantly applying mental strategies. We assessed mood with the Multidimensional Mood State Questionnaire (dimensions: good/bad, GB; awake/tired, AT; and calm/nervous, CN), and subjective arousal with Self-Assessment Manikins (SAM).

Results: We found significantly higher subjective arousal after neurofeedback phases in the experimental condition as compared to the control condition [t(26.6) = −2.216, 95% CI [−2.188,−0.083], p = 0.035; t(27.9) = −3.252, 95% CI [−2.685,−0.609], p = 0.003], but no significant differences between the conditions regarding mood [GB: b = 0.4, 95% CI [−0.67, 1.47], p = 0.467; AT: b = 0.769, 95% CI [−0.319, 1.857], p = 0.177; CN: b = 0.5, 95% CI [−0.53, 1.53], p = 0.352]. In both conditions, there was significantly worse and more tired mood after the fMRI session as
compared to before [GB: b = −0.77, 95% CI [−1.31, 0.23], p = 0.009; AT: b = −0.652, 95% CI [−1.116, −0.187], p = 0.01].

**Conclusions:** Findings indicate that rtfMRInf led to higher arousal, which may counteract the aim to reduce stress responses. Whether the multitasking situation has triggered this neurofeedback-related arousal – and how to circumvent it – asks for further study.

**Trial registration:** NCT01921088, ClinicalTrials.gov, 13th August 2013.

**Keywords**
arousal, dual task, multitasking, functional magnetic resonance imaging, mood, neurofeedback, psychological stress
Introduction

Stress is ubiquitous in life. Getting rid of it is neither realistic nor desirable, as Hans Selye pointed out: ‘complete freedom from stress is death’[1], p 137. However, accepting stress as part of our lives does not mean we are all heading towards long-term detrimental consequences of this inevitability. We can influence how a stressor affects us in the long run. For this, let us begin with how we react to a stressor: The stress response is manifested in physiological and psychological aspects and leads to specific behavior. Physiologically, stress for example increases blood pressure, heart rate, and specific brain activity, and triggers a cascade of endocrine activity which ends in glucocorticoid release[2]. Psychologically, stress leads to focused attention and increased arousal and alertness, and shows up in behavioral measures such as self-reported questionnaires[3].

This response ensures survival in the presence of a life-threatening stressor, or at least increases the chance of survival. In everyday life in a modern society, however, the response might overshoot, given the nature of the stressors we are facing. And accumulated or chronic stress can lead to dire mental and physical health issues: Stress-related mental disorders like depression and anxiety, which are globally a main source of adult disability[4], for example, or hypertension[5], which is the top modifiable risk factor for mortality[6].

The effects of stress and related disorders burden industrialized countries increasingly. Associated costs have been estimated to 20 billion euros per year, in the European Union alone[7]. It is also a topic which affects most of us eventually: 49% of people in a survey in the United States had a major stressful event in a one year time-window[8]. When asked how often they experience stress in their daily life, 44% of respondents said “frequently” and 35% “sometimes” to a Gallup poll conducted in the US in 2017[9].

Finding ways to better deal with stressors in the short term might spare us of these long-term consequences and lessen the burden on the individual and society. In the definition of psychological stress by Lazarus and Folkman, our own appraisal of a situation and our coping abilities take a major role: “a relationship with the environment that the person appraises as significant for his or her well-being and in which the demands tax or exceed available coping resources”[10], p 61.

Cognitive and behavioral techniques used in occupation-specific stress management programs are also known to psychotherapeutic practice. Intervention programs for stress management often include education about and practice of time-management and coping skills, psychoeducation, relaxation techniques (e.g. Jacobson’s progressive muscle relaxation, controlled breathing, hypnosis), mindfulness-based stress reduction, exercise, leveraging social support, or training in specific job-related skills to prevent or prepare for common stressors[11-14]. Yoga and meditation-based therapies have both also been associated with mood changes in people with the stress-related disorders depression and anxiety[15]. The mechanism behind these changes might act via the biological stress system[16].

However, while there is support for classical stress management interventions[17], as with any treatment it is likely that a proportion of affected people do not respond to a given intervention. In such cases, innovative neuroscientifically-informed interventions might help. Such interventions are based on neuroscientific knowledge about the stress response and can be coupled with personal neural activity in reaction to a stressor. Due to that, they are likely to help participants work deliberately on their individual stress response.

Gaining deliberate control over specific and individual brain activity (and thus indirectly over mental processes) is a key aim of neurofeedback, an approach that has over the last decades been applied to modulate a wide set of mental processes and to improve symptoms to specific mental disorders[18]. Neurofeedback describes the paradigm to feed back a signal reflecting a person’s own brain activity so that the person can use information contained in the signal to better modulate their brain activity[19]. An advantage of fMRI is that it allows to work with a spatially circumscribed region of the brain. Thus, one can use this approach to target various brain areas associated with different mental processes, in real-time functional magnetic resonance imaging neurofeedback (rtfMRInf). rtfMRInf has been applied to modulate diverse mental processes, including pain[20], mood[21], and many others[22]. Whether the procedure could also be used to modulate the central and peripheral stress response has to the best of our knowledge not yet been investigated.

Within a larger rtfMRInf study, we assessed self-reported mood and arousal measures and tested whether these differed between the participants who had received real feedback from their own brain’s activity, as compared to those who received sham feedback (the recorded brain activity of another participant). Participants thereby used one out of four different mental strategies to help them reduce their stress response. The use of these strategies has shown to improve mood when trained once a day for 13 consecutive days, using smartphone-based instructions[23]. Here, we aimed to elucidate whether rtfMRInf aiming at modulating the central and peripheral stress response is related to changes in mood and subjective arousal. While we addressed the primary and secondary outcomes of the study, namely physiological components of stress (brain activity and blood pressure) and adverse events, in another manuscript (Belardi, Lee, Kim, Stalujanis, Jung, Oh, Yoo, Pruesnner, Tegethoff, Meinlschmidt; unpublished study), here we focus on additional outcomes, namely self-reported psychological measures of mood and arousal.

We assumed that effects of rtfMRInf on mood and arousal may show up in one of two potential directions: Neurofeedback may lead to i) improved mood and lower arousal (in line with its aim to reduce the stress response); or ii) worse mood and higher arousal (in line with increased mental workload based on the multitasking situation going along with rtfMRInf). The second direction might be due to our experiment requiring the participants to multitask: at the same time monitoring the feedback signal; applying a specific mental strategy; and conducting the Stroop task. Current research in the field of
multitasking generally reports lower task performance when the task is performed in a multitasking setting, as compared to when it is done as a single task\(^4\) and higher arousal for complex and multitasking situations\(^5,6\).

**Methods**

**Participants**

We recruited 31 subjects and analyzed data of 30 of these (with mood and arousal data lacking from one subject, due to no-show for the main experiment). They were all male students at Korea University, Seoul, Republic of Korea. We allocated 17 of them to the experimental condition and 13 to the control condition, based on a predefined block-randomized scheme (blocks of 8, 10, and 12 with random order). The mean age was 24.6 years (SD=2.1) and 24.5 years (SD=2.4) in the experimental and the control condition, respectively with a mean of education of 14.9 years (SD=1.4) and 15.3 years (1.3), respectively. There were no significant differences between the conditions for these baseline characteristics [age: t(28) = -0.0585, p = 0.954; education years: t(28) = 0.718, p = 0.479].

We based the sample size on previous studies which had shown large effect sizes for rtfMRInf\(^7,8\). Using power analysis, we estimated that with 14 subjects in each group we could detect effects of \(d = 1.0\) with sufficient power (\(1 - \beta > .80\); given \(\alpha = 0.05\), one-sided). Recruitment was stopped after the intended 30 subjects had participated in the experiment.

A researcher in Switzerland who was not directly involved in conducting the experiment and who had no contact to the participants, generated the randomized allocation sequence. MATLAB was used for the randomization, whose underlying random number generator uses the Mersenne Twister algorithm by default\(^8\). This researcher ensured that the allocation sequence was concealed from those who recruited and assigned the participants. Participants were assigned to a condition according to the allocation sequence in the order in which they were included in the study and only after a final decision about inclusion was made, to support concealment. Researchers at Korea University did the enrollment and then assigned participants to the conditions.

**Sampling procedure**

We recruited participants via ads on the university’s website and a bulletin board on campus. Using the following inclusion and exclusion criteria, we checked all interested students and decided about their eligibility for the study. Inclusion criteria were being i) male, ii) between 18 and 65 years of age, iii) right-handed, iv) familiar with using a smartphone, to take part in the ambulatory training, v) having sufficient English language skills to follow the written instructions in the experiment, vi) no indication of color-blindness, vii) no history of cardiovascular or neurological diseases, and viii) no history of a severe mental disorder. After finishing the whole study procedure, we paid each participant 60,000 KRW to compensate for time and effort related to study participation.

The institutional board of Korea University approved the study and all participants gave written informed consent (approval number: KU-IRB-10-38-A-2(E-A-1)(E-A-1)(E-A-3)).

**Materials**

We used established tools to assess mood and arousal, applied a well-known cognitive task as a stressor in the fMRI experiment, and instructed our participants in four mental strategies, aimed at reducing their stress response during the experiment.

To assess mood, we used the English version of the established multidimensional mood state questionnaire (MDMQ) (original in German “Mehrdimensionaler Befindlichkeitsfragebogen (MDBF)”), which has good psychometric properties\(^9,10\). The questionnaire measures current mood on three dimensions: good to bad, awake to tired, and calm to nervous. Individual values on each dimension range from 4 to 24 and higher values represent more positive affect, feeling more awake, and calmer, respectively.

We assessed subjective arousal with a non-verbal pictorial rating scale to assess valence, arousal, and dominance, on a 9-point Likert Scale, called Self-Assessment Manikin (SAM)\(^1\). The arousal rating was labeled with “At the moment, I’m feeling…” and went from “very calm” to “very aroused” in addition to the original pictures. The SAM is an established tool which is used extensively in research. We were only interested in the arousal dimension, because it is clearly linked to a psychological stress response and, thus, have not analyzed the other dimensions of the SAM.

During the rtfMRInf experiment, we induced acute stress using a cognitive task which had previously been used for this purpose and shown to elicit a cardiovascular and neural stress response: the Stroop color-word interference task\(^2\), adapted for the use in fMRI experiments and to be more challenging due to implemented adaptive time constraints\(^3,4\). We instructed the participants in four mental strategies: Body attention, contemplative repetition (mantra), emotional imagery, and facial expression (make different emotional faces). More details about these strategies and the exact instructions (text and video clips) were published elsewhere\(^5\).

**Overall study procedure**

We laid out the whole study as a randomised parallel-group study with rtfMRI neurofeedback as the experimental condition and sham-feedback as the control condition. Participants were blinded about their allocation, while experimenters and those analyzing the data were not. We registered the study before starting recruitment (ClinicalTrials.gov, identifier NCT01921088). Over the course of the study, participants visited the laboratory three times and conducted 13 days of smartphone-based ambulatory mental training between the two main experimental visits. We conducted the study at the Korea University, Seoul, Republic of Korea (RRID:SCR_004095) between August and October of 2013.
Initially, we screened all interested students in a short telephone interview to check for their history of diseases and mental disorders described in the exclusion criteria above. On a first visit to the laboratory, we then further checked their eligibility based on all additional inclusion and exclusion criteria with a set of questionnaires. There, we also instructed participants about the whole study and the four mental strategies.

After deciding upon the final inclusion in the study, participants then visited the laboratory two more times for the main rtfMRInf experiments. These two visits were 14 days apart and during the 13 days in-between, participants took part in a smartphone-based ambulatory mental training, where they applied the mental strategies they had already used during the experiment in short daily sessions. They were guided through the training with video clips and questionnaires on their smartphones.

We here report data from the first laboratory visit (screening day) and the first experiment day and will, thus, refer to the latter day simply as “experiment day”. More details on the procedures, especially regarding the ambulatory training, have been reported elsewhere.

Experimental procedure
With regard to the first experiment day, the experimental procedure contained the following phases: Structural scans, where the previously defined broad regions of interest for the neurofeedback training were localized; Functional localizer phase, where participants did the Stroop task and the individual regions of interest could be pinpointed, to ensure participants get a feedback signal from areas active during the task; resting phase; Neurofeedback-only (i.e., without Stroop) phase, where participants first had to apply the four learned mental strategies in turn, and then continue using the strategy which worked best for them, and also do neurofeedback; resting phase; phase with additional structural scans; Neurofeedback with Stroop, where subjects used both, the mental strategies and the neurofeedback signal to actively modulate their brain activity associated with the stroop; resting phase; Stroop-only phase, where subjects only used the mental strategies to reduce their stress response; resting phase.

The Stroop task runs were made up of 8 blocks each; congruent and incongruent trials were alternated. During the neurofeedback phases, we presented the feedback signal continuously on one side of the screen and (if applicable) the Stroop task on the other side. We assessed current mood (MDMQ questionnaire) once before and once after the whole fMRI experiment and during the 13 days in-between, participants took part in a smartphone-based ambulatory mental training, where they applied the mental strategies they had already used during the experiment in short daily sessions. They were guided through the training with video clips and questionnaires on their smartphones.

Neurofeedback
We defined a set of regions of interest (ROIs) encompassing the left and right anterior cingulate cortex (ACC) and insular cortex (IC), based on previously found brain activity associated with the Stroop task. Within this set, a more precise individual ROI was localized during the functional localizer phase for each participant. The recorded and processed brain activity of the individual ROIs was fed back to the participants in near real-time. Participants saw the feedback signal abstracted as a white, moving thermometer-like bar on a black background, which went up and down depending on the signal strength, indicating the divergence from the baseline activity level. We instructed them to modulate the activity of the ROIs using this information, by applying the mental strategies they had learned and the information from the feedback signal. Sham-feedback for the control condition was the recording of the feedback signal from another participant.

We acquired the MRI data with a 3T Siemens Tim Trio scanner with a 12-channel head coil (Erlangen, Germany). To measure the BOLD signal, we applied a standard gradient-echo EPI pulse sequence, using the following specifications for rtfMRInf: repetition time (TR) = 1500 ms, echo time = 25 ms, field of view 240*240 mm, matrix size 64*64, voxel size = 3,75*3,75*5 mm, flip angle 90, and 30 interleaved slices with 5mm thickness at approximately 30 oblique to the AC-PC line without a gap. We calculated individual ROIs for each participant during the functional localizer phase as follows: EPI preprocessing (head motion correction for six parameters, spatial smoothing with an 8 mm full-width at half maximum Gaussian kernel); estimation of beta-value maps for each incongruent and congruent Stroop trial via general linear model (GLM) implemented in SPM to get a contrast map for “incongruent > congruent Stroop trials”; calculating the neurofeedback signal then from the intersection map between ROIs from the GLM and the predefined set of ROIs.

To calculate the neurofeedback signal, we first removed possible artifacts from the raw BOLD signal of the individual ROIs, applying a bandpass-filter (0.008 - 0.1 Hz) using a third-order elliptic digital filter to avoid low-frequency linear drift. Next, we linearly detrended the median BOLD signals within each of the ROIs as well as the whole-brain area. We then averaged the values between the 10th and 30th percentile during the cross-fixation period, using this as the baseline BOLD intensity (for ROI and whole-brain area). Percentage signal change (PSC) of the ROI relative to the whole-brain area were then estimated voxel-wise, by subtracting the estimated whole-brain PSC from the ROI PSC. This PSC difference was used as the neurofeedback signal. Finally, we averaged the signal over the last three TR periods in order to reduce potential high-frequency fluctuations occurring due to cardiac-and respiratory-related activity.

Data analysis
All offline data analysis was conducted using the software package R (version 3.5.1 and later; RRID:SCR_001905) and specific further packages for R as follows: “lme4” (RRID: SCR_015654) and “lmeTest” (RRID:SCR_015656), to conduct the mixed effects models, “dplyr” (RRID:SCR_016708) and “tidyverse” (RRID:SCR_017102), for data preparation, and “ggplot2” (RRID:SCR_014601) and “ggpubr” to create and export data visualizations. For online fMRI data preparation, analysis, feedback signal calculation, and neurofeedback presentation...
as described above, we used MATLAB (The MathWorks, Inc., Natick, MA, USA; RRID:SCR_001622) with SPM8 (RRID:SCR_007037).

To take into account the longitudinal nature of the mood data (two measurements, before and after the fMRI session), we used linear mixed effects models. Our models included the following factors: fixed effects Time (prescan, postscan), Condition (experimental, control), and the interaction Time*Condition, and random intercept for each participant. We estimated three models, one for each of the mood dimensions as dependent variable. Together with beta values, we report 95% confidence intervals of two-sided tests using an alpha-level of 0.05 to determine statistical significance.

Results

Participant flow
The flow of participants, from enrollment to allocation and analysis, is given in Figure 1 in a flow diagram consistent with the Consolidated Standards of Reporting Trials (CONSORT).

Mood
Mood scores were assessed with the MDMQ once before and once after the fMRI session. Results of individual mixed models for the three mood dimensions are presented in Table 1, and descriptive statistics can be found in the interaction plots in Figure 2. In the mood dimensions good/bad and awake/tired, we saw lower values after the fMRI session than before in both conditions, leading to a significant main effect of time. In the calm/nervous dimension, a slight drop in values was present only in the experimental condition, but neither the time effect nor the interaction with condition was significant in this model. None of these models, to determine effects on mood, showed a significant main effect for condition or an interaction between time and condition.

Arousal
We calculated Welch two sample t-tests for unequal variances to assess differences in SAM arousal values between the experimental and control conditions. To account for heteroscedasticity, these tests model different variances for both levels of the factor condition. Descriptive values for subjective arousal can be found in Figure 3. In the “Neurofeedback-only” phase, we found SAM arousal to be significantly higher for participants in the experimental condition \[t(26.6)=-2.216, \text{95%CI}[-2.188, -0.083], p=0.035, \text{(two-tailed test)}\], as compared to those in the control condition. In the phase “Neurofeedback with Stroop”, SAM arousal was also significantly higher in the experimental condition \[t(27.9)=-3.252, \text{95%CI}[-2.685, -0.609], p=0.003, \text{(two-tailed test)}\]. This difference was not present in

---

**Figure 1. CONSORT flow diagram.**
### Table 1. Mixed models results of mood data.

#### Mood good/bad

<table>
<thead>
<tr>
<th>Predictors</th>
<th>Estimates</th>
<th>95%CI</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>15.7</td>
<td>14.63</td>
<td>16.76</td>
</tr>
<tr>
<td>Time (after fMRI-session)</td>
<td>-0.77</td>
<td>-1.31</td>
<td>-0.23</td>
</tr>
<tr>
<td>Condition (control)</td>
<td>0.4</td>
<td>-0.67</td>
<td>1.47</td>
</tr>
<tr>
<td>Time*Condition</td>
<td>0.17</td>
<td>-0.37</td>
<td>0.71</td>
</tr>
</tbody>
</table>

#### Random Effects

<table>
<thead>
<tr>
<th>Marginal R² / Conditional R²</th>
<th>σ²</th>
<th>ICC subject</th>
<th>Observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.072 / 0.632</td>
<td>4.28</td>
<td>6.5</td>
<td>59</td>
</tr>
</tbody>
</table>

#### Mood awake/tired

<table>
<thead>
<tr>
<th>Predictors</th>
<th>Estimates</th>
<th>95%CI</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>14.769</td>
<td>13.681</td>
<td>15.857</td>
</tr>
<tr>
<td>Time (after fMRI-session)</td>
<td>-0.652</td>
<td>-1.116</td>
<td>-0.187</td>
</tr>
<tr>
<td>Condition (control)</td>
<td>0.769</td>
<td>-0.319</td>
<td>1.857</td>
</tr>
<tr>
<td>Time*Condition</td>
<td>0.113</td>
<td>-0.351</td>
<td>0.577</td>
</tr>
</tbody>
</table>

#### Random Effects

<table>
<thead>
<tr>
<th>Marginal R² / Conditional R²</th>
<th>σ²</th>
<th>ICC subject</th>
<th>Observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.090 / 0.720</td>
<td>3.31</td>
<td>7.43</td>
<td>60</td>
</tr>
</tbody>
</table>

#### Mood calm/nervous

<table>
<thead>
<tr>
<th>Predictors</th>
<th>Estimates</th>
<th>95%CI</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>15.35</td>
<td>14.32</td>
<td>16.38</td>
</tr>
<tr>
<td>Time (after fMRI-session)</td>
<td>-0.28</td>
<td>-0.83</td>
<td>0.27</td>
</tr>
<tr>
<td>Condition (control)</td>
<td>0.5</td>
<td>-0.53</td>
<td>1.53</td>
</tr>
<tr>
<td>Time*Condition</td>
<td>0.28</td>
<td>-0.27</td>
<td>0.83</td>
</tr>
</tbody>
</table>

#### Random Effects

<table>
<thead>
<tr>
<th>Marginal R² / Conditional R²</th>
<th>σ²</th>
<th>ICC subject</th>
<th>Observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.039 / 0.573</td>
<td>4.62</td>
<td>5.78</td>
<td>59</td>
</tr>
</tbody>
</table>

*Note.* $\sigma^2$ = within-group variance, $\tau_{oo}$ = between-group variance, ICC = intraclass correlation coefficient, CI = confidence interval. Factor predictors were coded using effect/deviant coding to increase interpretability of the fixed effects. Comparison from the mean intercept of the factor to the level names in parentheses for each factor. In the good/bad and calm/nervous models, there was one missing value each.
Figure 2. Mood values reported before and after fMRI session for experimental and control condition. Lighter empty symbols in the background are individual data points (jittered on the x-axis to avoid overplotting), while filled symbols with error bars are condition group values (means and standard errors).

Figure 3. Subjective arousal by condition and experiment phase. Combination of dot plot with individual data points and a boxplot for each condition group. The boxplot's lower and upper hinges mark the 25th and 75th percentiles. Whiskers extend from the hinges to the largest/smallest value up to 1.5 * the inter-quartile range. Black X shapes mark the condition group means. There was one missing value in the experimental condition during the functional localizer phase.
with the help of neurofeedback. This could be linked to reduce the physiological and psychological stress response better mood after the training in accordance with the study aim. We observed no statistically significant mood changes associated with neurofeedback, but participants in both conditions reported worse mood and being more tired after the fMRI session as compared to before.

These findings pose several questions: First, why did subjective arousal rise, contrary to our goal to reduce stress with our intervention? Arousal rose for experimental condition participants but not for those in the sham-feedback control. This might have been due to increased mental load in subjects who got real feedback compared to those who got sham-feedback, because those receiving sham-feedback might have (consciously or unconsciously) realized that the shown signal was not contingent with their brain activity. They might then have given less attention to this signal and the neurofeedback training and could focus better on other task(s) (applying mental strategies and solving the Stroop task). Arousal levels might, thus, here be an indicator for multitasking and increased mental load instead of an effect of the Stroop-induced stress. In this sense, the multitasking aspect of the experiment may have itself become a stressor, because it increased the cognitive demand of the participant.

One possible explanation supporting the idea of increased cognitive demand in the experimental condition is the multitasking situation, present in the experiment. Participants had to do several tasks simultaneously. This might have lead to increased mental load in subjects who got real feedback compared to those who got sham-feedback, because those receiving sham-feedback might have (consciously or unconsciously) realized that the shown signal was not contingent with their brain activity. They might then have given less attention to this signal and the neurofeedback training and could focus better on other task(s) (applying mental strategies and solving the Stroop task). Arousal levels might, thus, here be an indicator for multitasking and increased mental load instead of an effect of the Stroop-induced stress. In this sense, the multitasking aspect of the experiment may have itself become a stressor, because it increased the cognitive demand of the participant.

The second set of questions concerning our results is: Why could we not observe any statistically significant mood changes related to the intervention and what could explain the increased tiredness and worse mood after the fMRI session?

We observed no statistically significant mood changes associated with neurofeedback, even though we would have expected better mood after the training in accordance with the study aim to reduce the physiological and psychological stress response with the help of neurofeedback. This could be linked to one limitation in the experimental design, namely that we measured mood only completely before and after the fMRI session. The observed mood effects can, thus, primarily be interpreted in relation to the participant’s experience over the whole session and unfortunately they can not be matched to putative mood changes related to single experiment phases. Interestingly, subjective arousal, which was measured directly after each experiment phase during the fMRI session, did show differences between the conditions. We can, thus, assume that the sampling rate of mood might not have been fine-grained enough to pick up differences between the conditions and was only able to represent the overall experiment effects on all participants.

Regarding tiredness, fatigue due to the experiment and cognitive demand is expected. A one hour fMRI session is tiring and such overall effects might overshadow the miniscule differences between conditions due to the manipulation (real vs. sham feedback). Especially also since the neurofeedback manipulation was only present in some phases of the experiment. The lower awake/tired mood values after the experiment are, thus, not surprising. Participants may become tired after a demanding experiment in an fMRI scanner where they have to repeatedly solve a monotonous cognitive task. Furthermore, we expected our participants to relax and calm down. Thus, their indication of being more tired can be interpreted in line with what they actually did.

To explain the decrease in the good/bad mood dimension, we can look at the individual items in the questionnaire that made up this dimension: Subjects rated to what degree they felt uncomfortable, content, discontent, good, bad, happy, unhappy, great, superb, and wonderful. For example, it is unlikely to feel more comfortable and content after the experiment, given that lying in an fMRI scanner can be somewhat uncomfortable, and considering that participants were challenged with a cognitive task with adaptive difficulty, ensuring that they did not perform too well. Even if participants could modulate their immediate stress response, the overall mood change from before to after the experiment towards a worse mood might, thus, be explainable.

We also did not target to specifically change mood with our intervention. In comparison to another rtfMRI study, which did exactly that\(^1\), we used different target brain regions. Where these researchers targeted brain regions that most highly reflected activity differences in response to positive vs. neutral images, we focused on regions associated with our stressor task. Our modulated regions were thus less likely to be directly involved in supporting positive mood and we could only expect a potential side-effect in the mood due to the down-regulation of the stress response.

Future studies should aim to overcome the above-mentioned limitations, including limited time resolution in mood assessment: They could profit from using a shorter mood assessment instrument that can be applied in higher frequency. One example is a visual analog scale (VAS) to rate perceived stress, which could be implemented during an fMRI experiment and which allows
to sample rapidly and repeatedly, yet with lower precision as compared to the MDMQ. Longer questionnaires like the applied MDMQ interrupt the experiment for a longer time and are thus not ideal to be applied in higher frequency.

Broader implications for future research include a notion that the multitasking situation during an experiment can itself influence the measured values. While it is important to make the best of experimental time for economic reasons and to avoid prolonging an experiment unnecessarily, overloading an experiment might result in unexpected and intertwined effects. In our case, the multitasking present during the experiment might have led to our finding of increased arousal connected to the neurofeedback intervention. Even though challenging, future rtfMRInf studies on stress should try to prevent multitasking situations as good as possible. One could also more explicitly look at this multitasking aspect and conduct experiments to elucidate this component in the context of rtfMRInf research.

We had set out asking whether rtfMRInf to modulate the stress response would influence participants’ subjective perception of mood and arousal. The mood effects reflected the overall experimental experience due to sampling only before and after the fMRI session, and probably reflected rather general fatigue due to the cognitively demanding experiment than specific neurofeedback effects. To the best of our knowledge, we are the first to report a phenomenon of neurofeedback-related arousal: With regard to arousal, our findings are in line with the assumption that the multitasking nature of conducting neurofeedback during a stress task may have increased acute stress perceived by our participants, being in contrast to short-term neurofeedback effects on reduced subjective indicators of stress. Future studies should take into account multitasking situations in the experimental design, and further elucidate the neurofeedback-related arousal phenomenon, especially in the context of stress.

Data availability
Underlying data
Full underlying (non-aggregated) data cannot be made publically available since the ethics approval of this study does not cover openly publishing non-aggregated data.

In order to access this data, it must be requested from the corresponding author. Data requestors will have to provide: i) written description and legally binding confirmation that their data use is within the scope of the study; ii) detailed written description and legally binding confirmation of their actions to be taken to protect the data (e.g., with regard to transfer, storage, back-up, destruction, misuse, and use by other parties), as legally required and to current national and international standards (data protection concept); and iii) legally binding and written confirmation and description that their use of this data is in line with all applicable national and international laws (e.g., the General Data Protection Regulation of the EU).

Reporting guidelines

The completed CONSORT checklist is available under the terms of the Creative Commons Attribution 4.0 International license (CC-BY 4.0).

Grant information
This study was funded by the Swiss National Science Foundation, SNSF (project no. PZ00P1_137023 to MT). Further, GM, JL, and MT received funding from the National Research Foundation of Korea (NRF) within the Global Research Network Program (under project no. 2013S1A2A2035364). JL received funding in form of a NRF grant, MSIP of Korea (NRF-2016M3C7A1914450), and in part in form of a National Research Council of Science Technology (NST) grant by the Korean government (MSIT) [No. CAP-18-01-KIST]. GM received funding from the Stanley Thomas Johnson Stiftung & Gottfried und Julia Bangerter-Rhyner-Stiftung under projects no. PC_28/17 and PC_05/18, from the Swiss Cancer League (Krebsliga Schweiz) under project no. KLS-4304-08-2017, from the Research Foundation of the International Psychoanalytic University (IPU) Berlin, from Gesundheitsförderung Schweiz under project no. 18.191, and from the Swiss National Science Foundation (SNSF) under project no. 100014_135328.

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Acknowledgements
We thank Peter J. Gianaros, Ph.D. from the Department of Psychology, University of Pittsburgh, Pittsburgh, PA, USA, for his valuable contributions to this study by supporting us with the paradigm central to our experiment and by sharing his insights during the design phase.

References
The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias
- You can publish traditional articles, null/negative results, case reports, data notes and more
- The peer review process is transparent and collaborative
- Your article is indexed in PubMed after passing peer review
- Dedicated customer support at every stage

For pre-submission enquiries, contact research@f1000.com