Similarities and distinctions between cortical neural substrates that underlie generation of malevolent creative ideas

https://doi.org/10.1523/ENEURO.0127-23.2023

Cite as: eNeuro 2023; 10.1523/ENEURO.0127-23.2023

Received: 20 April 2023
Revised: 30 August 2023
Accepted: 1 September 2023

This Early Release article has been peer-reviewed and accepted, but has not been through the composition and copyediting processes. The final version may differ slightly in style or formatting and will contain links to any extended data.

Alerts: Sign up at www.eneuro.org/alerts to receive customized email alerts when the fully formatted version of this article is published.

Copyright © 2023 Qiao et al.
This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International license, which permits unrestricted use, distribution and reproduction in any medium provided that the original work is properly attributed.
Running head: Neural correlates that underlie various malevolent creativity

Similarities and distinctions between cortical neural substrates that underlie generation of malevolent creative ideas

Xinuo Qiao \textsuperscript{a,1}, Kelong Lu \textsuperscript{b,1}, Qiang Yun \textsuperscript{a}, Ning Hao \textsuperscript{b,*}

\textsuperscript{a} Shanghai Key Laboratory of Mental Health and Psychological Crisis Intervention, School of Psychology and Cognitive Science, East China Normal University, Shanghai, China

\textsuperscript{b} School of Mental Health, Wenzhou Medical University, Wenzhou, Zhejiang, China

\textsuperscript{*} Address correspondence to: Ning Hao, Ph.D. Shanghai Key Laboratory of Mental Health and Psychological Crisis Intervention, School of Psychology and Cognitive Science, East China Normal University, No. 3663, North Zhong Shan Road, Shanghai 200062, China. Tel: +86 21 62232909. Fax: +86 21 62232909. E-mail: nhao@psy.ecnu.edu.cn

\textsuperscript{1} Xinuo Qiao and Kelong Lu contributed equally to this paper and should both be considered as first authors.
ORCID

Xinuo Qiao: 0000-0002-3875-3556
Kelong Lu: 0000-0002-5881-2437
Qiang Yun: 0000-0002-0313-9038
Ning Hao: 0000-0001-6784-6228

Number of Figures: 4
Number of Tables: 4
Number of words for Abstract: 261
Number of words for Introduction: 1752
Number of words for Discussion: 2045

Conflict of Interest
The authors declare that they have no competing interests.

Acknowledgements
This work was supported by the National Natural Science Foundation of China [31971002] to NH.
Abstract

Creativity can be driven by negative intentions, and this is called malevolent creativity (MC). It is a type of creativity that serves antisocial purposes and deliberately leads to harmful or immoral results. A possible classification indicates that there are three kinds of MC in daily life: hurting people, lying, and playing tricks. This study aimed to explore similar and distinct neural substrates underlying these different kinds of MC idea generation. The participants were asked to perform different MC tasks, and their neural responses were recorded using a functional near-infrared spectroscopy device. The findings revealed that most regions within the prefrontal and temporal lobes (e.g., the right dorsolateral prefrontal cortex, rDLPFC, and right angular gyrus) were involved in the three MC tasks. However, the right frontopolar cortex (rFPC) was more activated and less coupled with the rDLPFC and right precuneus during the lying task than during the other tasks. Thus, rFPC may play an important role in constructing novel lies. In the lying task, individuals were more selfish and less compassionate. In the playing tricks and hurting people tasks, there was less neural coupling between the rDLPFC and the left inferior frontal gyrus/right inferior parietal lobule than that in the lying task. This may imply that selfish motivation is released when individuals try to ignore victims’ distress or generate aggressive tricks in hurting people or playing tricks tasks. These findings indicate that the three kinds of MC idea generation involve common cortical regions related to creative idea generation and moral judgment, whereas differences in cortical responses exist due to their unique features.

Keywords: creativity, malevolent creativity, fNIRS, functional connectivity
Significant Statement

Malevolent creativity (MC) is a type of creativity that serves antisocial purposes and deliberately leads to harmful or immoral results. A possible classification indicates that there are three kinds of MC in daily life: hurting people, lying, and playing tricks. The present study, for the first time, explored the similarities and distinctions between cortical neural substrates that underlie generation of malevolent creative ideas. The findings indicate that the three kinds of MC idea generation involve common cortical regions related to creative idea generation and moral judgment, whereas differences in cortical responses exist due to their unique features. This study provided support for the MC three-dimensional framework (i.e., hurting people, lying, and playing tricks), and preliminarily revealed the underlying neural substrates of various MC idea generation, which helps to develop effective techniques or approaches to manage and prevent malevolent creative behaviors.
1. Introduction

Creativity is the ability to produce novel and useful work that usually contributes to positive outcomes (Runco & Acar, 2012; Sternberg & Lubart, 1996), but it can also be driven by evil intentions, resulting in negative outcomes. For example, an employee may tell a creative lie to steal credit from their colleagues. This creativity is termed malevolent creativity (MC), serves antisocial purposes, and deliberately leads to harmful or immoral results (Cropley et al., 2010; Cropley et al., 2008; Harris et al., 2013; Qiao et al., 2022).

According to the definition of MC, malevolent creative behaviors have negative consequences on individuals and society. The randomness, surprise, and rule-breaking features of creativity render malevolent behaviors unpredictable (Gill et al., 2013; Gino & Wiltermuth, 2014; Wang, 2019). Moreover, compared with uncreative unethical behaviors, creative ones are considered less immoral; thus, they are punished less severely (Wiltermuth et al., 2017). This implies that MC increases the success rate of malevolent behavior and makes it easier for implementers to escape punishment. Although MC is usually marked by creative criminal or terrorism activities (e.g., Cropley et al., 2008; Gill et al., 2013), it can be observed in the general population (e.g., Hao et al., 2020; Perchtold-Stefan et al., 2020a), such as framing other colleagues creatively to avoid punishment. These MC behaviors may not be considered crimes but still hurt others’ personal interests. Most people are more likely to encounter the outcomes of MC from ordinary people than from criminals or terrorists. Additionally, non-criminal malevolent creative behavior may escalate into criminal malevolent creative behavior in the future. Therefore, based on the findings from studies on criminals or terrorism, researchers have tried to understand MC in the general population. Several factors that are positively related to MC have been explored, including individual aggression (Lee &
Dow, 2011; Harris & Reiter-Palmon, 2015), antagonism (Perchtold-Stefan et al., 2020a), lower emotional intelligence (Harris et al., 2013), the Dark Triad (Gao et al., 2022b; Jonason et al., 2017), approach motivation (Hao et al., 2020), and anger emotional state (Cheng et al., 2021a; Cheng et al., 2021b). While, it would be preliminary and important to explore the underlying mechanisms (including neural substrates) of MC idea generation, which helps to develop effective techniques or approaches to manage and prevent malevolent creative behaviors, as a result prevents the negative effects of malevolent creativity on society.

To date, several studies have directly investigated the neural correlates of idea generation in MC. In these studies, the participants were required to generate creative solutions that caused damage. Qiao et al. (2022) reported that the middle occipital gyrus is less involved (i.e., weaker activation and less coupling with other regions) in the generation of malevolent creative ideas, which might reflect the inhibition of moral judgment. These results suggest that a decline in moral criteria supports the emergence of immoral ideas. Another study observed decreased efficiency within the default networks (i.e., the superior frontal gyrus and middle temporal gyrus) in MC idea generation, which might imply reduced social cognition toward victims (Gao et al., 2022a). In addition, lower functional connectivity strength in the postcentral gyrus (POG) may be related to weakened emotional perception (Gao et al., 2022a), while malevolent ideation may be weakened after enhancing activity in the right POG (Gao et al., 2023). Perchtold-Stefan et al. (2023a) used an EEG device to investigate task-related power changes in the alpha band during the generation of malevolent ideas. The results showed that individuals with higher MC performance had increased prefrontal alpha power in the early stage (indicating conceptual expansion from prosocial to antisocial) and increased temporal alpha power in the middle stage (indicating inhibition of dominant
but normal ideas; Perchtold-Stefan et al., 2023a). Furthermore, researchers have noticed gender differences in the neural substrates underlying MC. Although females and males have similar levels of MC performance, females tend to generate malevolent creative ideas through semantic memory retrieval and the recombination of social information, whereas males rely more on motor-related imagery (Perchtold-Stefan et al., 2023b).

However, MC is not a one-dimensional framework. Researchers have proposed a possible classification of MC behaviors in daily life, namely hurting people, lying, and playing tricks (Hao et al., 2016). These MC behaviors refer to the generation of creative solutions that can cause damage, but the way in which they cause damage is different. (1) Hurting people indicates obtaining an unfair advantage by harming others physically or mentally (e.g., attacking enemies with new weapons). (2) Lying is concerned with telling a lie, concealing, and framing (e.g., designing a novel scam). (3) Playing tricks is used to design a creative prank or practical joke (e.g., setting up a novel trick).

According to this definition, hurting people is more dangerous than the other two kinds of MC (Hao et al., 2016). While, lying and playing tricks can cause substantial financial or reputation losses. Given that hurting people has been the main focus of the aforementioned neuroscience MC studies, to further understand the neural portrait of MC, the neural substrates (brain activation and cross-regional neural coupling) that underlie the generation of various malevolent creative ideas deserve exploration. Therefore, this study aimed to reveal the similarities and distinctions between the neural substrates underlying idea generation in the three kinds of MC.

Although different kinds of MC vary greatly, they all involve generating creative ideas. Thus, neural responses in the creativity-related regions may be observed in all kinds of MC. According to previous studies, the executive control network (ECN; e.g., the middle frontal gyrus) and default
network (DMN; e.g., the right temporal-parietal junction [rTPJ] and precuneus) play important roles in the creative process (Beaty et al., 2016; Beaty et al., 2017; Fink et al., 2010; Wu et al., 2015; Zhu et al., 2017). Specifically, the executive control network is related to working memory and task-set switching, which are necessary for creative idea generation (Beaty et al., 2016; Zhu et al., 2017), whereas the default network is involved in memory retrieval (e.g., rTPJ) and information integration (e.g., angular gyrus [AG]) in creativity (Fink et al., 2010; Zhu et al., 2017). Furthermore, the executive control network cooperates with the default network by inhibiting prominent responses elicited by memory retrieval (Beaty et al., 2017). Moreover, MC is characterized by damaging and breaking social norms; thus, all kinds of MC are related to antisocial behavior and morality. Consistent with this, a previous study found that individuals who were less concerned about moral values tended to have higher MC levels (Kapoor & Kaufman, 2022a). At the neural level, antisocial behavior is related to the dysfunction of the right temporal lobe and reduction of the gray matter volume in the left frontopolar cortex (lFPC) and right postcentral gyrus (rPOG) (Aoki et al., 2014; Bertsch et al., 2013; Fumagalli & Priori, 2012). Individuals tend to break reciprocal norms when the right dorsolateral prefrontal cortex (rDLPFC) is inhibited (Knoch et al., 2006). Accordingly, these findings suggest that all kinds of malevolent creative idea generation may be involved in the neural responses of the bilateral PFC and right temporal lobe.

Regardless of their similarities, different kinds of MC differ in specific features. Hurting people is directly related to aggression (Hao et al., 2016). Studies showed that increased physical aggression is accompanied by decreased neural activation or PFC impairment, which may reflect a deficiency in inhibition (Nelson & Trainor, 2007; Soloff et al., 2003). Aggressive individuals exhibit abnormal neural activation in the right precuneus (Wong et al., 2018). Moreover, structural and functional
reductions in the AG and superior temporal gyrus are associated with increased hurting behaviors (Glenn & Raine, 2014; Raine, 2019). In addition to the PFC, the temporal lobe also plays an important role in generating ideas that hurt people. Lying is characterized by creative and dishonest behavior. Previous research revealed that the frontal (i.e., DLPFC) and temporal regions are involved in dishonest behavior (Lisofsky et al., 2014; Ofen et al., 2017; Sun et al., 2015; Yin et al., 2016). This suggests that cognitive control and perspective-taking play essential roles in cheating others (Lisofsky et al., 2014; Ofen et al., 2017). Additionally, the right frontopolar cortex (rFPC) is involved in well-rehearsed lies (Abe, 2011; Ganis et al., 2003). Given that creative lies are usually more subtle and complex, rFPC may be involved in weaving malevolent creative lies. Playing tricks is typically used to create fun. Studies demonstrated increased activity in the left inferior frontal gyrus (lIFG) during humor detection (Moran et al., 2004). The theory-of-mind network (i.e., the middle temporal gyrus) is more involved in point-to-other jokes (Feng et al., 2014). Brain regions associated with humor and jokes may overlap with those associated with generating creative tricks.

This study aimed to reveal similar and distinct neural correlates of the three kinds of MC idea generation using a one-way factorial design, with TASK (tasks of hurting people, lying, and playing tricks) as the within-subject factor. The participants were required to complete three MC tasks: (1) hurting people, (2) lying, and (3) playing tricks. They were instructed to generate different creative malevolent ideas using the corresponding kind of MC. For example, the participants were only allowed to generate creative methods to cause physical or mental damage during the hurting people task. Neural activity during each task was recorded using functional near-infrared spectroscopy (fNIRS), which has several advantages, such as (1) higher tolerance of body movement (Schecklmann et al., 2010), (2) higher ecological validity, and (3) greater economy. The common
cerebral regions involved in all MC tasks were identified by comparing the neural activation and cross-regional neural coupling of all MC tasks with those at baseline. Subsequently, neural activation and coupling were compared among all MC kinds to reveal their specific neural correlates. The hypotheses were as follows: (H1) Neural responses in brain regions such as the IFPC, middle frontal gyrus, rPOG, and right MOG (rMOG) would be observed in the three kinds of MC. (H2a) The hurting people task is more associated with the right precuneus, AG, and superior temporal gyrus. (H2b) The lying task is more associated with the DLPFC and rFPC. (H2c) The playing tricks task is more associated with the lIFG and middle temporal gyrus. Moreover, previous studies found that the Dark Triad of personality, moral personality, general creativity potential, and MC potential can affect MC performance (Gao et al., 2022b; Jonason et al., 2017; Kapoor & Kaufman, 2022b; Perchtold-Stefan et al., 2023b). Therefore, we also measured these indices using a series of scales and tested whether these variables were distinctively related to the neural substrates of the different kinds of MC idea generation.

2. Method

2.1. Participants

A total of 40 participants (mean age = 21.30 ± 2.23 years; 32 females; college students) participated in the study (two participants were excluded due to poor signal, see details in the “fNIRS data pre-processing” section). A priori power analysis showed that the sufficient sample size to obtain reliable results was 36 (1-β = .90, α = .05, effect size f = .25; Chow et al., 2017), which is comparable to previous neuroscientific creativity studies (Takeuchi et al., 2019; Tempest & Radel, 2019). All the participants were right-handed and had normal or corrected-to-normal visual acuity. None of the patients had any history of mental or neurological illness. Written informed consent was
obtained from all participants before the experiment. Each participant received compensation of ¥55.

The study procedure was approved by the University Committee on Human Research Protection of the ******** (Code ********). To avoid misrepresentation, the sample in the current study was collected independently and did not overlap with any published studies.

2.2. Experimental procedure

Upon arrival, each participant was asked to sit in front of a laptop table. The experimenter introduced the experimental procedure to the participants, prepared the fNIRS device, and told them to be ready. The initial scan session lasted for 30 s and served as the baseline, during which the participants were required to close their eyes, remain still, and relax. The task session consisted of 3 blocks (10 trials per block). The participants were required to solve 10 hurting people, lying, or playing tricks tasks respectively during each block. The sequence of the blocks was counterbalanced among the participants.

Each trial successively consisted of an 8 sec fixation session, a 10 sec task-reading session, a 20 sec thinking session, and a 12 sec reporting session. The participants verbally reported their most creative responses in the reporting session (only one response was allowed for each trial). Two jitters (blank screen, 2–6 s) were set between the latter three sessions (Fig. 1A).

Immediately after each block, the participants rated their anxiety, pleasure, and tasks’ benevolence and malevolence on a 7-point Likert scale ranging from 1 (not at all) to 7 (strongly). The ratings of tasks’ malevolence and benevolence were designed to verify the effectiveness of the tasks (see details in the “Experimental tasks” section). There was a 1 min resting session after each block. Eventually, the participants were asked to complete a series of scales right after they finished the whole experiment (see details in the “Post-experimental scales” section).
2.3. Experimental tasks

The hurting people tasks, lying tasks and playing tricks tasks were adopted from the previous study. It showed that there were no significant differences in the aspects of difficulty, fatigue, involvement and familiarity among these three MC tasks (Cheng et al., 2021b).

**Hurting People Task (HPT).** This task was adapted from a realistic presented problem (RPP), which requires individuals to generate a creative solution for an open-ended realistic problem (Agnoli et al., 2016; Runco et al., 2016; Xue et al., 2018). Each HPT problem was expressed using words limited to 38–42 Chinese characters. The participants were explicitly required to solve the problem by causing physical or mental damage in novel ways (e.g., Hong encounters a tennis master in the final game, which makes it difficult for her to win. Please generate a novel way for Hong to injure her opponent “accidently”). It should be noted that the participants could generate several ideas; however, they only needed to report the idea they considered the most creative. The participants were asked to rate the intelligibility, malevolence, and benevolence of the 10 HPT problems on a Likert scale ranging from 1 (very low) to 7 (very high). These 10 HPT problems showed that: (1) intelligibility was all above 6, (2) malevolence was all above 5, (3) benevolence was all below 3.

**Lying Task (LT).** The procedure for developing this task was similar to that for the hurting people task. The participants were required to solve the problem by lying, concealing, and framing in novel ways (e.g., Ming has stolen public donations for private use but does not want to be caught. Please generate a novel way for Ming to conceal this thing). The 10 LT problems’ intelligibility, malevolence, and benevolence were above 6, above 5, and below 3 respectively.
Playing Tricks Task (PTT). The procedure for developing this task was similar to that for the hurting people task. The participants were required to solve the problem using a creative prank or practical joke (e.g., A head teacher who is very strict toward Jie intends to host an open class. Please generate a novel way for Jie to ruin the open class without being caught). The intelligibility, malevolence, and benevolence of the 10 PTT problems were above 6, above 5, and below 3 respectively.

2.4. Assessment of performance during the hurting people, lying, and playing tricks tasks

Performance during the hurting people, lying, and playing tricks tasks was assessed by evaluating the originality and harmfulness of the ideas (Runco & Okuda, 1991; Gao et al., 2022a; Perchtold-Stefan et al., 2020a, 2022; Qiao et al., 2022). As only one idea was allowed for each task, idea fluency or uniqueness was not assessed (Lu et al., 2019a). Thus, a widely used subjective method was employed to assess originality and harmfulness of the ideas (Beaty et al., 2018; Lu et al., 2019b; Xue et al., 2018). Five trained raters (with at least three years of experience in creativity research) independently scored originality and harmfulness of the ideas on a 5-point Likert scale (originality: 1 = not original at all, 5 = highly original; harmfulness: 1 = not harmful at all, 5 = highly harmful). The inter-rater agreement was satisfactory (internal consistency coefficient [ICC]; originality: ICC [hurting people task] = .79, ICC [lying task] = .70, ICC [playing tricks task] = .72; harmfulness: ICC [hurting people task] = .80, ICC [lying task] = .74, ICC [playing tricks task] = .70). The originality and harmfulness scores for each idea were calculated by averaging the scores of all the raters. The final scores of the participants were calculated by averaging the scores for all the ideas.

2.5. Post-experimental tests

The Runco Ideational Behavior Scale (RIBS; e.g., I have a lot of ideas about stories and poetry;
Runco et al., 2016) was used to evaluate the participants’ general creativity potential. The RIBS contains 19 items rated on a 5-point Likert scale ranging from 0 (never) to 4 (approximately every day). The internal consistency reliability was satisfactory (Cronbach’s α = .88). The Malevolent Creativity Behavior Scale (MCBS; e.g., I have thought of some new methods to punish people who do something wrong; Hao et al., 2016) was used to evaluate the participants’ MC potential. The MCBS contains 13 items rated on a 5-point Likert scale ranging from 1 (never) to 5 (always). The internal consistency reliability was satisfactory (Cronbach’s α = .84). The Honesty-Humility Inventory (Ho-Hu; e.g., If I knew I would never get caught, I would want to steal a million yuan; Lee & Ashton, 2004) was used to evaluate the participants’ moral personalities. The Ho-Hu contains 20 items rated on a 5-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). The internal consistency reliability was satisfactory (Cronbach’s α = .78). The Chinese version of the Dirty Dozen (DD12; e.g., I am used to getting my own way by manipulating others; Geng et al., 2015) was used to evaluate the participants’ Dark Triad personality traits. The DD12 contains 12 items rated on a 7-point Likert scale ranging from 1 (strongly disagree) to 7 (strongly agree). The internal consistency reliability of the DD12 was unsatisfactory in this study (Cronbach’s α = .68); therefore, this scale was not included in the subsequent analysis.

2.6. fNIRS data acquisition

An fNIRS device (ETG-7100, Hitachi Medical Corporation, Japan) was used to record the concentrations of oxyhemoglobin (HbO) and deoxyhemoglobin (HbR) during the experiment. The absorption of near-infrared light (wavelengths of 695 and 830 nm) was assessed at a sampling rate of 10 Hz. The bilateral PFC and right temporal region were selected as ROIs in this study. A 3×5 probe set (eight emitters and seven detectors, 3 cm optode separation, forming 22 measurement channels;
Fig. 1B) was placed on the bilateral PFC, and one 4×4 probe set (eight emitters and eight detectors, 3
cm optode separation, forming 24 measurement channels; Fig. 1C) was placed on the right temporal
region. According to the 10-20 system, the lowest row of the 3×5 probe set was aligned with the
Fp1–Fp2 line, and optode “A” was positioned on the frontal pole midline point (Fpz; Sai et al., 2014).
In the meantime, the middle probe column of the 3×5 probe set was aligned along the sagittal
reference curve (Fig. 1B). The 4×4 probe set was aligned with the sagittal reference curve, and
optode “B” was positioned on P6 (Fig. 1C). The virtual registration method was used to establish the
correspondence between channels and measurement points (brain regions) in the cortex (Singh et al.,
2005; Tsuzuki et al., 2007).

2.7. fNIRS data pre-processing
The principal component spatial filter algorithm was used to remove the global components of
the raw fNIRS data, whereas the correlation-based signal improvement method was used to remove
motion artifacts (Cui et al., 2012; Pan et al., 2018; Fig. 1D). Considering that the HbO and HbR
signals were negatively correlated after using the correlation-based signal improvement method (the
corrected HbR is solely the corrected HbO multiplied by a negative coefficient; Cui et al., 2010), the
subsequent analysis mainly focused on HbO signals.
Channels with poor signals were determined by visually checking the NIRS time-course plot.
The channel with a considerably higher variance than other channels of the same participant was
identified as a poor channel (e.g., the variances of normal channels were 0.5–0.8, but those of poor
channels were 10–30). Participants with more than 11 poor channels (25% of the total channels)
were excluded from the subsequent analysis.

2.8. Data analysis of neural activation
The NIRS Statistical Parametric Mapping (NIRS_SPM) package was used to estimate individual neural activation in this study (Jang et al., 2009; Ye et al., 2009). The hrf low-pass filtering and wavelet minimum description length detrending algorithms were selected as required by the NIRS_SPM. Neural activation during the entire thinking stage (0-20 s) was estimated using a general linear model (Fig. 1D). First, a reference wave was set for each channel for all conditions (baseline, hurting people task, lying task, and playing tricks task) to represent the theoretical variations in Hbo signals induced by the experimental stimulus. Then, a regression analysis containing theoretical HbO signal variations and real HbO signal variations during the task period (baseline and thinking stage of hurting people, lying, and playing tricks tasks) was performed for each channel. Beta (β) values, which indicate variation in neural activation, were obtained as regression coefficients for all channels under different conditions. Furthermore, the beta increment was calculated by subtracting the baseline beta value from the thinking-session beta value for the three tasks.

Primarily, one-sample t-tests using 0 as the test value (baseline) were performed on the beta increments of each task. All *p*-values were corrected using the false discovery rate correction method (FDR; corrected alpha level = 0.05). The channels that were significant in all three tasks were marked as common activated or common deactivated.

Next, the beta increments were Z-scores transformed channel by channel across the participants. A total of 46 one-way repeated measures ANOVAs, using TASK as the within-subject factor, were performed on the beta increments. All *p*-values were corrected using the FDR method (corrected alpha level, 0.05). Subsequent post hoc tests were corrected using Bonferroni correction when necessary.

A nonparametric permutation test was performed on significant channels (Maris & Oostenveld,
The procedure was as follows: (1) within each channel, we randomly sampled the data from three MC tasks to reconstruct three new subsets; (2) one-way repeated measures ANOVAs using three new subsets as within-subject factors were conducted to calculate a new $F$ value; (3) we repeated (1) and (2) 5000 times to establish a distribution of $F$ values (i.e., 5000 one-way repeated measures ANOVAs on 5000 three subsets); (4) if the actual $F$ value was higher than 95% of the $F$ values in the distribution (permutation threshold $p < 0.05$), differences among the three tasks were considered significant.

Linear regression was performed to quantify the relationship between beta increments from all the channels and task performance. Before the formal linear regression procedure, several Pearson correlation analyses were performed on the behavioral performance and beta increments of each channel. Only the channels that were significantly correlated with behavioral performance ($p < 0.05$, uncorrected) were introduced into the subsequent analysis as predictors (Xie et al., 2022). In this study, linear regression was performed on every behavioral performance (with beta increments as predictors). In addition, the relationships between the post-experimental scales and beta increments were examined using linear regression (with post-experimental scales as dependent variables and beta increments as predictors).

### 2.9. Data analysis of neural coupling

Neural coupling was introduced to measure the functional connectivity between different cerebral regions in this study. Cross-correlations were used to assess the NC between time-series Hbo concentrations from different cerebral regions during different MC tasks, which were suitable for evaluating the covariation of two signals over time (Liu et al., 2021; Ward et al., 2018; Fig. 1D). The NC values (i.e., cross-correlation coefficients) were converted using Fisher’s z-transformation.
The NC increment was calculated by subtracting the baseline NC value from the NC values of the three tasks.

Initially, there were 2,116 channel (CH) combinations (46 × 46 CHs). After excluding 1,081 redundant CH combinations (equal CH combinations or CH combinations of a single CH, i.e., CH1-CH1), only 1,035 valid CH combinations were finally entered into the subsequent analyses.

Likewise, several one-sample t-tests using 0 as the test value (baseline) were performed on the NC increments of each task. All p-values were corrected using the FDR method (corrected alpha level, 0.05). The NC increments of CH combinations that were significantly higher than baseline in all the three tasks were marked as the common increased NC increments. Next, several one-way repeated measures ANOVAs, using TASK as the within-subject factor, were performed on the NC increment for each CH combination. The p-values were corrected using the FDR method across all CH combinations (1,035 combinations; corrected alpha level, 0.05). Subsequent post hoc tests were corrected using Bonferroni correction when necessary. A nonparametric permutation test was performed on significant NC increments to validate the results (Maris & Oostenveld, 2007; Zheng et al., 2020).

Linear regression analysis, similar to the beta increment, was performed to quantify the relationship between the NC increments of all the CH combinations and task performance. Considering that there were 1,035 CH combinations, only the CH combinations that were significantly correlated with behavioral performance (p < .005, uncorrected) were introduced into the subsequent linear regression as predictors (Xie et al., 2022). The remaining procedure was the same as that used for neural activation (see details in the “Data analysis of neural activation” section). Additionally, linear regression with the same procedure was used to examine the relationship...
between post-experimental scales and NC increments.

### 2.10. Code accessibility

The code used in this study is available from the corresponding author upon request.

### 3. Results

#### 3.1 Behavioral indices

One-way repeated measures ANOVAs, using TASK as the within-subject factor, were performed for anxiety, pleasure, benevolence, and malevolence. The results showed that the main effect of TASK on anxiety was significant ($F(2, 78) = 4.17, p = .019, \eta_p^2 = .10$). The post hoc test (Bonferroni-corrected) revealed that the participants experienced significantly more anxiety during the hurting people task than during the playing tricks task ($p = .043$). The results also showed a significant main effect of TASK on pleasure ($F(2, 78) = 5.20, p = .008, \eta_p^2 = .12$). The post hoc test revealed that the participants significantly experienced more pleasure during the playing tricks task than during the lying task ($p = .004$). Detailed descriptive data are presented in Table 1.

The main effect of TASK on benevolence was significant ($F(2, 78) = 5.17, p = .008, \eta_p^2 = .12$). The post hoc test (Bonferroni-corrected) revealed that the playing tricks task contained significantly more benevolence than the hurting people task ($p = .005$). A one-sample $t$-test (test value = 3) revealed that the benevolence of the hurting people ($t(39) = -14.70, p < .001$, Cohen’s $d = 2.32$), lying ($t(39) = -9.40, p < .001$, Cohen’s $d = 1.49$), and playing tricks tasks ($t(39) = -10.07, p < .001$, Cohen’s $d = 1.59$) were significantly lower than 3. The main effect of TASK was also significant for malevolence ($F(2, 78) = 13.42, p < .001, \eta_p^2 = .26$). The post hoc test (Bonferroni-corrected) revealed that the hurting people task contained more malevolence than the lying ($p < .001$) and playing tricks tasks ($p < .001$). A one-sample $t$-test (test value = 5) revealed that the malevolence of the hurting...
people task ($t(39) = 23.24$, $p < .001$, Cohen’s $d = 3.68$), lying task ($t(39) = 6.83$, $p < .001$, Cohen’s $d = 4.08$), and playing tricks task ($t(39) = 4.54$, $p < .001$, Cohen’s $d = 1.08$) was significantly higher than 5. These results partially confirm the effectiveness of the three tasks. Detailed descriptive data are presented in Table 1.

### 3.2. Beta increments

One-sample $t$-tests using 0 as the test value (baseline) were performed on the beta increments of each task among all the CHs (see details in Table 3). After FDR correction ($p < .05$, a total of 46 channels), the results revealed that the bilateral FPC (CH2, CH3, and CH4), right middle frontal gyrus (CH9), rAG (CH39), and right precuneus (CH43 and CH46) were more activated than baseline; the rDLPFC (CH16), right middle temporal gyrus (CH23), right superior temporal gyrus (CH27), rPOG (CH30), and right supramarginal gyrus (rSMG; CH34) were less activated than baseline.

One-way repeated measures ANOVAs, using TASK as the within-subject factor, were performed on the beta increments of all CHs. Individual data with poor signals were excluded from each CH. After FDR correction ($p < .05$, a total of 46 channels), the results revealed that the main effects of TASK were significant on the beta increment at CH3 ($F(2,76) = 8.40, p = .023, \eta^2_p = .18$) (Fig. 2B) and CH4 ($F(2,76) = 7.47, p = .025, \eta^2_p = .16$) (Fig. 2C). Both CH3 and CH4 are located in the rFPC. Post hoc tests (Bonferroni-corrected) showed that the beta increment at CH3 was significantly higher during the lying task ($M = .15, SD = 1.08$) than during the hurting people task ($M = -.05, SD = .89, p = .009$) and playing tricks task ($M = -.10, SD = 1.03, p = .002$). Moreover, the beta increment at CH4 was significantly higher during the lying task ($M = .26, SD = .94$) than during the hurting people ($M = -.13, SD = .89, p = .003$) and playing tricks task ($M = -.13, SD = 1.13, p = .006$). The permutation test revealed that the $F$ values of these results were larger than the 95% empirical distribution (Fig.
3.3. Neural couplings

One-sample \(t\)-tests, using 0 as the test value (baseline), were performed on the NC increments of each task among all CH combinations (see details in Table 4). After FDR correction (\(p < .05\), a total of 1035 valid channels), the following NC increments were more coupled than baseline: \(lFPC\) and left DLPFC (\(lDLPFC; \text{CH1-CH5}\)), right fusiform and rMOG (\(CH24-CH25\)), rAG and rMOG (\(CH32-CH36\)), rSMG and rAG (\(CH35-CH38, \text{CH38-CH39, and CH39-CH42}\)), and rPOG and right precentral gyrus (\(CH41-CH44 \text{ and CH44-CH45}\)).

Several one-way repeated measures ANOVAs, using TASK as the within-subject factor, were performed for the NC increments of the valid CH combinations. Individual data with poor signals were excluded from each CH combination. After FDR correction (\(p < .05\); a total of 1,035 valid CH combinations), the results revealed that the main effects of TASK on the NC increments were significant for the following CH combinations: \(CH8-CH21, CH12-CH46, CH13-CH14, CH17-CH38, \text{and CH27-CH31}\). Because \(CH27\) and \(CH31\) were in the same region, this result was not included in further analyses (Fig. 3). These CH combinations refer to the rFPC, rDLPFC, right precuneus, IDLPFC, and right inferior parietal lobule (rIPL).

Specifically, post hoc tests (Bonferroni-corrected) revealed that the NC increment of \(CH8-CH21\) (rFPC-rDLPFC) was significantly lower during the lying task than during the hurting people (\(p < .001\)) and playing tricks task (\(p = .004\)) (Fig. 3A). The NC increment of \(CH12-CH46\) (rFPC-right precuneus) was significantly lower during the lying task than during the hurting people (\(p = .009\)) and playing tricks task (\(p < .001\)) (Fig. 3B). The NC increment of \(CH13-CH14\) (rDLPFC-lIFG) was significantly lower during the playing tricks task than during the hurting people (\(p = .019\)) and lying
task ($p < .001$) (Fig. 3C). The NC increment of CH17-CH38 (rDLPFC-rIPL) was significantly lower during the hurting people task than during the lying ($p = .002$) and playing tricks task ($p = .002$) (Fig. 3D). Detailed descriptive data are listed in Table 2. The permutation test revealed that the $F$ values of these results were greater than 95% of the empirical distribution.

Considering that the levels of anxiety and pleasure were significantly different across the three tasks, the linear mixed model was introduced to control for the potential effect of anxiety and pleasure on the results of beta increments and neural couplings. The ANOVA results remained significant even after controlling for anxiety and pleasure. This suggests that these neural results are independent of anxiety and pleasure.

### 3.4. Linear regression analysis

#### 3.4.1. The originality and harmfulness of MC tasks

Linear regression analysis was used to quantify the relationship between neural results (beta increments and neural couplings) under different conditions and task performances (i.e., originality and harmfulness of the hurting people, lying, and playing tricks tasks). The results revealed that for the hurting people task, harmfulness was negatively predicted by the activation of rPOG (CH30; $b = -19.70$, 95%CI = [-.15, 35.58], $F = 6.03$, $p = .005$) and originality was negatively predicted by the neural coupling of lDLPFC-rPOG (CH5-CH35; $b = -1.65$, 95%CI = [-2.78, - .53], $F = 10.09$, $p < .001$). For the lying task, originality was positively predicted by the neural coupling between rFPC and rAG (CH12-CH35; $b = 1.60$, 95%CI = [.47, 2.73], $F = 5.50$, $p < .001$). For the playing tricks task, harmfulness was negatively predicted by the activation of rDLPFC (CH21; $b = -20.94$, 95%CI = [-40.85, -1.03], $F = 4.53$, $p = .040$) and originality was negatively predicted by the neural coupling between the right superior temporal gyrus and rSMG (CH26-34; $b = -.80$, 95%CI = [-1.51,
3.4.2. Post-experimental scales

Linear regression analysis was used to quantify the relationship between the neural activity and the scores of the post-experimental scales. The results revealed that the score of the DD12 (Dark Triad) was negatively predicted by the activation of the rIFG during the hurting people task (CH18; \( b = -42.95, 95\% CI = [-73.26, -12.63], F = 4.89, p = .002 \)). The score of the Ho-Hu (moral personality) was negatively predicted by the activation of the right fusiform gyrus during the hurting people (CH28; \( b = -85.67, 95\% CI = [-169.64, -1.70], F = 4.34, p = .001 \)) and lying tasks (CH28; \( b = -83.68, 95\% CI = [-169.64, -1.70], F = 5.60, p < .001 \)). The score of the RIBS (i.e., general creativity potential) was negatively predicted by the activation of the lDLPFC during the hurting people (CH5; \( b = -55.63, 95\% CI = [-109.31, -1.95], F = 8.36, p = .001 \)) and lying tasks (CH5; \( b = -58.01, 95\% CI = [-115.39, -1.62], F = 4.19, p = .048 \)) as well as the neural coupling of rFPC-lDLPFC during the playing tricks task (CH8-CH15; \( b = -3.35, 95\% CI = [-6.33, -0.37], F = 7.72, p = .001 \)). The score of the MCBS (i.e., MC potential) was negatively predicted by the neural coupling of rDLPFC-rPOG during the hurting people task (CH22-CH30; \( b = -1.24, 95\% CI = [-2.41, -0.07], F = 6.57, p < .001 \)), the score of the lying task was positively predicted by the neural coupling of lDLPFC-rFPC (CH5-CH8; \( b = 0.70, 95\% CI = [0.11, 1.28], F = 5.85, p < .001 \)), and the score of the playing tricks task was negatively predicted by the neural coupling of rMOG-rSMG (CH25-CH42; \( b = -0.63, 95\% CI = [-1.16, -0.09], F = 6.48, p = .001 \)).

4. Discussion

This study aimed to reveal similar and distinct neural correlates of the three kinds of malevolent creative idea generation (i.e., hurting people, lying, and playing tricks). The cerebral activity was
recorded using an fNIRS device. Several regions of the PFC and temporal lobe were involved in all kinds of MC idea generation, including the bilateral FPC, bilateral DLPFC, and right angular gyrus (rAG). For each kind, the rFPC during lying task was more activated and less coupled with the rDLPFC and right precuneus than baseline. The rDLPFC was less coupled with the IIFG/rIPL than baseline when playing tricks/hurting people (see summary figures, Fig. 4A-4D).

The results revealed that some regions (including those proposed in H1) were involved in all three kinds of MC. The bilateral FPC, right middle frontal gyrus, and right precuneus were more activated in malevolent creativity, whereas the neural coupling of the IFPC-IDLPFC was stronger than that at baseline. Increased activity in these regions is associated with general creativity (Chen et al., 2015; Green et al., 2015; Wu et al., 2015; Zhu et al., 2017). Moreover, the right middle temporal gyrus and right superior temporal gyrus were less activated, which may reflect higher openness and divergent thinking (Vartanian et al., 2008; Zabelina et al., 2019). These results indicate that individuals attempted to generate creative ideas during all tasks. Furthermore, the rDLPFC, rPOG, and right supramarginal gyrus (rSMG) were less activated than they were at baseline. Reduced activity in these regions involves selfish behavior, such as releasing self-interest motivation (Knoch et al., 2006), less empathy (Seehausen et al., 2016), and more emotional egocentricity (Silani et al., 2013). Additionally, the results showed that the activity of the rAG was higher than that at baseline and more coupled with the rMOG and rSMG. Neural coupling between the right fusiform gyrus and the rMOG, rPOG, and right precentral gyrus was stronger than that at baseline. Previous studies found that the rAG, right fusiform gyrus, and right precentral gyrus are associated with attention reorientation (Caspers et al., 2014; Hahn et al., 2006; Seghier et al., 2013). The activities of the MOG, rSMG, and rPOG are associated with prosocial behaviors, such as moral judgment (Cheng et al.,
2021), inhibiting emotional egocentricity (Silani et al., 2013), and empathy (Seehausen et al., 2016). These results suggest that individuals convert their attention to antisocial task requirements to avoid being affected by prosocial tendencies.

As predicted, we observed distinctions in neural responses among the three kinds of MC idea generation. Specifically, rFPC activation was stronger during the lying task than during the other tasks, which is consistent with H2b. A previous study showed that the bilateral frontal poles (especially the right frontal pole) are associated with theory of mind (ToM; Lewis et al., 2011; Reniers et al., 2014), which is closely related to lying or deception (Poletti et al., 2011; Sai et al., 2021; Walczyk & Cockrell, 2021; Walczyk et al., 2014). Liars rely on ToM to manipulate victims’ beliefs, predict their potential actions, and speculate about their thoughts (Walczyk & Cockrell, 2021; Walczyk et al., 2014). Moreover, rFPC is involved in well-rehearsed lies (Abe, 2011; Ganis et al., 2003). Research indicates that the frontopolar region is associated with high-level executive strategies for deception (Abe et al., 2006; Ganis et al., 2003; Lin et al., 2021). These findings emphasize the importance of rFPC in weaving well-rehearsed, convincing, and creative lies.

The NC increment in the rFPC-rDLPFC was significantly lower during the lying task than during the hurting people and playing tricks tasks. As stated above, rFPC is involved in well-rehearsed lies, and we observed an increase in beta increments during the lying task. Research showed that individuals’ honesty increases after neural augmentation of the rDLPFC using transcranial direct current stimulation (Maréchal et al., 2017). In this case, a decrease in the NC of the rFPC-rDLPFC may allow individuals to be dishonest and generate creative lies. In addition, we observed a decrease in the NC of the rFPC-right precuneus when compared to the hurting people and playing tricks tasks. The right precuneus is associated with compassion, a social emotion aroused by
personal loss and social deprivation (Immordino-Yang et al., 2009). Accordingly, such a decrease in
the rFPC-right precuneus may indicate that personal compassion toward victims is inhibited when
generating a creative lie. Linear regression analysis demonstrated that neural coupling between the
rFPC and rAG positively predicted the originality of the lying task. The rFPC is related to ToM
(Lewis et al., 2011; Reniers et al., 2014), whereas rAG is associated with understanding others’ states
(Lu et al., 2021; Schurz et al., 2017; Filmer et al., 2019). These mental processes play a pivotal role
in lying or deception (e.g., Walczyk & Cockrell, 2021). Thus, stronger coupling between the rFPC
and rAG may indicate that individuals try to know what others are thinking, which may contribute to
generating an original lie.

The results of the playing tricks task were partially consistent with H2c: the NC increment of the
rDLPFC-lIFG was lower during the playing tricks task than during the other tasks. Playing tricks
may be associated with dark humor, which is characterized by transgression from social norms and
moral systems (e.g., hurting or upsetting or demonstrating superiority over others; Aillaud & Piolat,
2012; Perchtold-Stefan et al., 2020b). In addition, dark humor is positively correlated with MC
(Perchtold-Stefan et al., 2020b). The rDLPFC is related to harnessing self-interest motivation (Knoch
et al., 2006), whereas the lIFG is associated with cognitive humor processing (Bekinschtein et al.,
2011; Samson et al., 2008; Vrticka et al., 2013). Thus, the lIFG may participate in the construction of
fun tricks. Meanwhile, individuals may need to be more “selfish” to obtain fun. Therefore, the
decrease in the NC increments of the rDLPFC-lIFG indicates that individuals focus on obtaining
self-interested pleasures when generating a “dark” trick to upset others. Additionally, activation of
the rDLPFC negatively predicts the harmfulness of playing tricks. In line with previous findings, this
result suggests that controlling selfish motivation causes low harmfulness in the playing tricks task.
(Knoch et al., 2006). The results also showed that higher coupling between the right superior temporal gyrus and the rSMG significantly negatively predicted the originality of the playing tricks task. The right superior temporal gyrus is activated during humor processing (Shibata et al., 2014). Increased activity of the rSMG is linked to reduced emotional egocentricity (Silani et al., 2013). Thus, awareness of a victim’s emotions may hinder the construction of a novel trick. Thus, stronger coupling between the right superior temporal gyrus and rSMG may indicate lower originality of the playing tricks task.

The NC increment of the rDLPFC-rIPL was lower than that of the other MC tasks. The rDLPFC is also involved in controlling selfish motivation (Knoch et al., 2006). A previous study showed that the rIPL participates in maintaining the self-other distinction (Uddin et al., 2006). Individuals experience more distress induced by others if the self-other distinction is inadequate (Krol & Bartz, 2021). Considering that the hurting people task requires participants to hurt the victims directly (e.g., physical damage), participants may be more likely to realize the victims’ negative feelings. Thus, this NC result implies that individuals may need to release selfish motivation and avoid being affected by victims’ emotions during the hurting people task, which further maximizes the performance of hurting people. The involvement of the right precuneus, AG, and superior temporal gyrus (proposed in H2a) in the hurting people task was not different from that in other tasks, which might be due to the uniqueness of suppressing victims’ emotions when hurting people. Moreover, the performance of the hurting people task was negatively predicted by the activation of the rPOG and neural coupling of the lDLPFC-rPOG. Previous studies demonstrated that the rPOG is associated with empathy (Seehausen et al., 2016) and emotional perception (Gao et al., 2022a). The prefrontal cortex plays a central role in appropriate moral behavior (including IDLPFC; Dashtestani et al., 2018; Forbes et al., 2021).
Therefore, low coupling between the lDLPFC and the rPOG indicates that individuals try to suppress the process of knowing the victim’s state and behave immorally, leading to more original and harmful ideas.

The results of linear regression analysis showed that neural activity in the three kinds of MC idea generation significantly predicted the level of Dark Triad, moral identity, and general and MC potential. Activation of the rIFG during the hurting people task negatively predicted the level of Dark Triad. Studies showed that the rIFG is involved in inhibition (Aron et al., 2003). Individuals with less inhibition toward immoral ideas during the hurting people task tended to have higher Dark Triad. The results revealed that moral identity was negatively predicted by the activity of the right fusiform gyrus during the hurting people and lying tasks. These results imply that individuals who actively reorient their attention toward malevolent task requirements may have low moral identity (Caspers et al., 2014). However, this effect was not significant in the playing tricks task, maybe because playing tricks was considered less morally related. Moreover, the results showed that activation and neural coupling within the prefrontal cortex were negatively related to general creativity potential in the three MC tasks. The activity of the prefrontal cortex is related to appropriate moral behavior (Dashtestani et al., 2018; Forbes et al., 2010), which is contrary to the requirements of the three MC tasks. Non-obedience of creative task requirements may indicate less general creativity potential. As for the hurting people and playing tricks tasks, the neural coupling of rDLPFC-rPOG and rMOG-rSMG negatively predicted hurting people and playing tricks potential. The rDLPFC and rMOG are associated with prosocial mental processes, such as controlling selfish motivation (Knoch et al., 2006) and moral judgment (Qiao et al., 2022). The activities of the rPOG and rSMG are involved in the emotional perception of victims (Gao et al., 2023; Silani et al., 2013). Thus, the
connectivity between these regions may weaken the generation of malevolent ideas. Another study found that neural coupling between the IDLPFC and rFPC positively predicted the lying potential. Previous studies showed that the rFPC is linked to well-rehearsed lies (Abe, 2011; Ganis et al., 2003), whereas the ECN (including the IDLPFC) is important in idea generation (Beaty et al., 2016; Zhu et al., 2017). Consequently, stronger neural coupling between the IDLPFC and rFPC could be an indicator of a higher lying potential.

In general, these findings contribute to better understanding the neural substrates underlying the generation of different creative malevolent ideas, which highlights the crucial role of the rFPC, rDLPFC, rIPL, and lIFG. Based on these findings, specific precautions should be taken for the different kinds of MC. For example, empathy interventions may be effective in preventing individuals from finding creative ways to hurt others. Moreover, the neural substrates of malevolent idea generation by criminals and terrorists should be further explored in the future. This may be helpful in predicting and preventing extremely harmful malevolent behaviors.

Several limitations should be mentioned: (1) fNIRS could only detect the outer cortex in the PFC and the right temporal and parietal regions. However, subcortical areas may also be involved in idea generation in different kinds of MC. Therefore, devices with higher spatial resolution (fMRI and MEG) should be used to explore the neural correlates of idea generation in different kinds of MC. (2) There was gender imbalance in this study sample. The effects of potential gender differences should be examined in the future. (3) This study focused only on MC in laboratory experiments. Whether the neural correlates of generating different malevolent creative ideas are dependent on the context (laboratory vs. real life) should be further examined. (4) It should be noted that participants completed post-experimental scales after three MC tasks. Engagement in malevolent idea generation
may influence participants’ subsequent evaluation of their behavior in daily life, which may affect the scores of scales. (5) Only limited variables (i.e., the Dark Triad of personality, moral personality, general and malevolent creativity potential) were measured in this study; other related variables (e.g., antagonism, state and trait anger, aggression, and emotion regulation) should be investigated in future studies to explore the relationship between these related variables and the neural substrates of different kinds of MC idea generation. Therefore, caution should be exercised when generalizing the findings of this study.

References


Lu K, Qiao X, Hao N (2019a) Praising or keeping silent on partner’s ideas: Leading brainstorming in


Perchtold-Stefan CM, Fink A, Rominger C, Papousek I (2020b) Motivational factors in the typical display of humor and creative potential: The case of malevolent creativity. Front


Perchtold-Stefan CM, Rominger C, Papousek I, Fink A (2023b) Women and men have a similar potential for malevolent creativity – but their underlying brain mechanisms are different. *Brain Res*, 1801: 148201.


Runco MA, Okuda SM (1991). The instructional enhancement of the flexibility and originality


Table 1

The results of one-way repeated measures ANOVAs with TASK as the within-subject factor on the anxiety, pleasure, benevolence and malevolence.

<table>
<thead>
<tr>
<th></th>
<th>HPT (M ± SD)</th>
<th>LYT (M ± SD)</th>
<th>PTT (M ± SD)</th>
<th>F(2, 78)</th>
<th>p</th>
<th>η²</th>
<th>Post-hoc</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anxiety</td>
<td>5.21 ± 1.44</td>
<td>4.97 ± 1.49</td>
<td>4.59 ± 1.58</td>
<td>4.17</td>
<td>.019</td>
<td>.10</td>
<td>HPT&gt;PTT*</td>
</tr>
<tr>
<td>Pleasure</td>
<td>2.28 ± 1.40</td>
<td>1.86 ± .76</td>
<td>2.46 ± 1.28</td>
<td>5.20</td>
<td>.008</td>
<td>.12</td>
<td>PTT&gt;LYT**</td>
</tr>
<tr>
<td>Benevolence</td>
<td>1.45 ± .67</td>
<td>1.70 ± .87</td>
<td>1.79 ± .76</td>
<td>5.17</td>
<td>.008</td>
<td>.12</td>
<td>PTT&gt;HPT**</td>
</tr>
<tr>
<td>Malevolence</td>
<td>6.68 ± .46</td>
<td>5.95 ± .88</td>
<td>5.84 ± 1.17</td>
<td>13.42</td>
<td>.000</td>
<td>.26</td>
<td>HPT&gt;LYT***</td>
</tr>
</tbody>
</table>

HPT>PTT***
Table 2

The results of one-way repeated measures ANOVAs with TASK as the within-subject factor on the NC increments.

<table>
<thead>
<tr>
<th>CH combination</th>
<th>HPT(\bar{M} \pm SD)</th>
<th>LYT(\bar{M} \pm SD)</th>
<th>PTT(\bar{M} \pm SD)</th>
<th>F</th>
<th>p(_{FDR})</th>
<th>(\eta^2)</th>
<th>Post-hoc</th>
</tr>
</thead>
<tbody>
<tr>
<td>CH8-CH21</td>
<td>.00 ± .86</td>
<td>-.53 ± .91</td>
<td>-.04 ± 1.11</td>
<td>9.49</td>
<td>.037</td>
<td>.20</td>
<td>LYT&lt;PTT***</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>LYT&lt;PTT**</td>
</tr>
<tr>
<td>CH12-CH46</td>
<td>-.20 ± .66</td>
<td>-.51 ± .67</td>
<td>.07 ± .95</td>
<td>10.56</td>
<td>.048</td>
<td>.22</td>
<td>LYT&lt;HPT**</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>LYT&lt;PTT***</td>
</tr>
<tr>
<td>CH13-CH14</td>
<td>.16 ± 1.36</td>
<td>.35 ± 1.21</td>
<td>-.25 ± 1.05</td>
<td>9.56</td>
<td>.043</td>
<td>.21</td>
<td>PTT&lt;HPT*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>PTT&lt;LYT***</td>
</tr>
<tr>
<td>CH17-CH38</td>
<td>-.37 ± .97</td>
<td>.33 ± .96</td>
<td>.24 ± 1.01</td>
<td>9.66</td>
<td>.048</td>
<td>.21</td>
<td>HPT&lt;LYT**</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>HPT&lt;PTT**</td>
</tr>
</tbody>
</table>
Table 3

Results of One-Sample T-test on beta Increment.

<table>
<thead>
<tr>
<th>Channel</th>
<th>t-values</th>
<th>Region</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>HPT</td>
<td>LYT</td>
</tr>
<tr>
<td>2</td>
<td>3.42**</td>
<td>4.16***</td>
</tr>
<tr>
<td>3</td>
<td>5.72***</td>
<td>5.91***</td>
</tr>
<tr>
<td>4</td>
<td>4.48***</td>
<td>6.84***</td>
</tr>
<tr>
<td>9</td>
<td>2.90*</td>
<td>3.10*</td>
</tr>
<tr>
<td>16</td>
<td>-2.65*</td>
<td>-2.60*</td>
</tr>
<tr>
<td>23</td>
<td>-3.06*</td>
<td>-3.91**</td>
</tr>
<tr>
<td>27</td>
<td>-5.59***</td>
<td>-3.11*</td>
</tr>
<tr>
<td>30</td>
<td>-3.66**</td>
<td>-4.11**</td>
</tr>
<tr>
<td>34</td>
<td>-3.46**</td>
<td>-4.27**</td>
</tr>
<tr>
<td>39</td>
<td>5.19***</td>
<td>5.00***</td>
</tr>
<tr>
<td>43</td>
<td>5.70***</td>
<td>5.73***</td>
</tr>
<tr>
<td>46</td>
<td>3.58**</td>
<td>3.95**</td>
</tr>
</tbody>
</table>
## Table 4

Results of One-Sample T-test on NC Increment.

<table>
<thead>
<tr>
<th>CH Combination</th>
<th>t-values</th>
<th>Region</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>HPT</td>
<td>LYT</td>
</tr>
<tr>
<td>CH1-CH5</td>
<td>3.74*</td>
<td>5.55***</td>
</tr>
<tr>
<td>CH24-CH25</td>
<td>3.50*</td>
<td>4.98**</td>
</tr>
<tr>
<td>CH32-CH36</td>
<td>3.91*</td>
<td>7.93***</td>
</tr>
<tr>
<td>CH35-CH38</td>
<td>4.00*</td>
<td>3.59*</td>
</tr>
<tr>
<td>CH38-CH39</td>
<td>4.18*</td>
<td>6.02***</td>
</tr>
<tr>
<td>CH39-CH42</td>
<td>5.60***</td>
<td>4.80**</td>
</tr>
<tr>
<td>CH41-CH44</td>
<td>4.23*</td>
<td>3.22*</td>
</tr>
<tr>
<td>CH44-CH45</td>
<td>7.19***</td>
<td>5.81***</td>
</tr>
</tbody>
</table>
Table Legends

Table 1
The post-hoc tests were Bonferroni corrected. HPT = hurting people task; LYT = lying task; PTT = playing trick task. *p < 0.05; **p < 0.01; ***p < 0.001.

Table 2
Only CH (channel) combinations that survived the FDR correction were listed below. The post-hoc tests were Bonferroni corrected. HPT = hurting people task; LYT = lying task; PTT = playing trick task. *p < 0.05; **p < 0.01; ***p < 0.001.

Table 3
Only channels that significant different from baseline in all tasks were listed below. The post-hoc tests were Bonferroni corrected. FPC = frontopolar cortex; MFG = middle frontal gyrus; DLPFC = dorsolateral prefrontal cortex; MTG = middle temporal gyrus; STG = superior temporal gyrus; POG = postcentral gyrus; SMG = supramarginal gyrus; AG = angular gyrus. *p < 0.05; **p < 0.01; ***p < 0.001.

Table 4
Only CH (channel) combinations that significant different from baseline in all tasks were listed below. The post-hoc tests were Bonferroni corrected. FPC = frontopolar cortex; DLPFC = dorsolateral prefrontal cortex; MOG = middle occipital gyrus; AG = angular gyrus; SMG = supramarginal gyrus; POG = postcentral gyrus; PRG = precentral gyrus. *p < 0.05; **p < 0.01; ***p < 0.001.
**Figure Legends**

**Figure 1**
Experimental design in the study. (A) Experimental procedure. Rest: 30-second resting state session; Rating: 30-second self-rating on anxiety, pleasure, benevolence and malevolence; The task sequence was counterbalanced among participants. (B) Optode probe set on the prefrontal cortex. (C) Optode probe set on the right temporal regions. (D) The data pre-processing and analysis procedure.

**Figure 2**
Results on beta increment (FDR-corrected). HP = hurting people; LY = lying; PT = playing trick. (A) The full views of main effect of TASK on the beta increment of all CHs. The red rectangle indicate that the CHs have significant main effect (CH3 and CH4). (B) The amplitude of beta increment of CH3. *p < 0.05, **p < 0.01, ***p < 0.001. (C) The amplitude of beta increment of CH4. *p < 0.05, **p < 0.01, ***p < 0.001. (D/E) The distribution of F values from permutation test on CH3/CH4. The arrows denote the location of actual F value.

**Figure 3**
Results on NC increment (FDR-corrected). HP = hurting people; LY = lying; PT = playing trick. (A-D) The amplitude of NC increments of CH8-CH21, CH12-CH46, CH13-CH14 and CH17-CH18. The distributions of F values from permutation test on these CH combinations. The arrows denote the locations of actual F values. *p < 0.05, **p < 0.01, ***p < 0.001.

**Figure 4**
Summary figure of results. DLPFC = dorsolateral prefrontal cortex; IPL = inferior parietal lobule; FPC = frontopolar cortex; IFG = inferior frontal gyrus; POG = postcentral gyrus; AG = angular gyrus; STG = superior temporal gyrus; SMG = supramarginal gyrus; (A) results of neural activation; the upward arrow means that the neural activation during certain task was significantly weaker than other tasks (e.g., the neural activation at right FPC during lying task was significant stronger than hurting people and playing trick tasks); (B) results of neural coupling; the downward arrow means that the neural coupling during certain task was significantly weaker than other tasks (e.g., the neural coupling between right DLPFC and right IPL during hurting people task was significantly weaker...
than during lying and playing trick tasks); (C) results of linear regression analysis on neural activation; the symbol (-) means negative prediction (e.g., the neural activation of right DLPFC during playing trick significantly negatively predict the performance of playing trick task); (D) results of linear regression analysis on neural coupling; the symbol (+) means positive prediction (e.g., the neural coupling between right FPC and right AG during lying significantly positively predict the performance of lying task).
A

CH3  CH4
Right frontopolar cortex

B

C

D

E

5% area

5% area
<table>
<thead>
<tr>
<th></th>
<th>Hurting people</th>
<th>Lying</th>
<th>Playing trick</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>A</strong> Neural activation relative to Baseline</td>
<td>/</td>
<td>Right PFC</td>
<td>/</td>
</tr>
<tr>
<td><strong>B</strong> Neural coupling relative to Baseline</td>
<td>Right DLPFC-Right IPL</td>
<td>Right FPC-Right precuneus Right FPC-Right DLPFC</td>
<td>Right DLPFC-Right IFG</td>
</tr>
<tr>
<td><strong>C</strong> Neural activation predicting task performance</td>
<td>Right POG (+) → Harmfulness</td>
<td>/</td>
<td>Right DLPFC (-) → Harmfulness</td>
</tr>
<tr>
<td><strong>D</strong> Neural coupling predicting task performance</td>
<td>Left DLPFC-Right POG (-) → Originality</td>
<td>Right FPC-Right AG (+) → Originality</td>
<td>Right STG-Right SMG (-) → Originality</td>
</tr>
</tbody>
</table>