Direct stimulation of anterior insula and ventromedial prefrontal cortex disrupts economic choices

Romane Cecchi, Antoine Collomb-Clerc, Inès Rachidi, Lorella Minotti, Philippe Kahane, Mathias Pessiglione, Julien Bastin

nature communications

Article

https://doi.org/10.1038/s41467-024-51822-8

Direct stimulation of anterior insula and ventromedial prefrontal cortex disrupts economic choices

Received: 18 December 2023

Accepted: 16 August 2024

Published online: 29 August 2024

Check for updates

Romane Cecchi[1,5,6] ✉, Antoine Collomb-Clerc[1,7], Inès Rachidi[1,2], Lorella Minotti[1,2], Philippe Kahane[1,2], Mathias Pessiglione[3,4] & Julien Bastin[1] ✉

Neural activity within the ventromedial prefrontal cortex (vmPFC) and anterior insula (aIns) is often associated with economic choices and confidence. However, it remains unclear whether these brain regions are causally related to these processes. To address this issue, we leveraged intracranial electrical stimulation (iES) data obtained from patients with epilepsy performing an economic choice task. Our results reveal opposite effects of stimulation on decision-making depending on its location along a dorso-ventral axis within each region. Specifically, stimulation of the ventral subregion within aIns reduces risk-taking by increasing participants' sensitivity to potential losses, whereas stimulation of the dorsal subregion of aIns and the ventral portion of the vmPFC increases risk-taking by reducing participants' sensitivity to losses. Moreover, stimulation of the aIns consistently decreases participants' confidence, regardless of its location within the aIns. These findings suggest the existence of functionally dissociated neural subregions and circuits causally involved in accepting or avoiding challenges.

Although it is widely accepted that risky decisions are influenced by the desirability of potential outcomes, the neural mechanisms underlying such choices remain poorly understood. Evidence from functional imaging and intracranial electrophysiology suggests that the ventromedial prefrontal cortex (vmPFC) and the anterior insular cortex (aIns) play a critical role in shaping these decisions. Accordingly, increased pre-stimulus neural activity in the vmPFC and aIns has been shown to promote and temper risk-taking, respectively, by overweighting the prospects of monetary gain or loss[1,2]. However, while the timing of these activities suggests a possible causal relationship between these brain regions and choice behavior, the correlational nature of the techniques used to date prevents conclusive establishment of this causal relationship.

In parallel, another line of research has shown that neural activity in both the vmPFC[3–6] and aIns[7–10] is related to participants' confidence in their own decisions. Moreover, confidence judgments appear, in turn, to be influenced by monetary prospects[11,12], suggesting that the neural codes for value and confidence may not be independent.

To date, causal studies investigating the involvement of the vmPFC and aIns in risky decision-making and confidence judgments have provided limited evidence due to several limitations. Firstly, the deep anatomical location of these brain regions has posed challenges in targeting these areas using non-invasive stimulation methods in

[1]Univ. Grenoble Alpes, Inserm, U1216, CHU Grenoble Alpes, Grenoble Institut Neurosciences, 38000 Grenoble, France. [2]Neurology Department, University Hospital of Grenoble, Grenoble, France. [3]Motivation, Brain and Behavior (MBB) team, Paris Brain Institute, Pitié-Salpêtrière Hospital, Paris, France. [4]Université de Paris, Paris, France. [5]Present address: Laboratoire de Neurosciences Cognitives et Computationnelles, Institut National de la Santé et de la Recherche Médicale, Paris, France. [6]Present address: Département d'Études Cognitives, École Normale Supérieure, Université Paris Sciences et Lettres, Paris, France. [7]Present address: NeuroX Institute and Brain Mind Institute, School of Life Sciences, Swiss Federal Institute of Technology (EPFL), Lausanne, Switzerland. ✉e-mail: romane.cecchi@gmail.com; julien.bastin@univ-grenoble-alpes.fr

humans[13–15]. Secondly, lesion studies have yielded inconsistent results regarding risk-taking behavior that may be attributed to variations in the scope and precise location of brain lesions[16–24]. Finally, intracranial electrical stimulation (iES) studies have rarely examined these regions[25,26] or have not combined iES with cognitive paradigms designed to tease apart precise decision-making processes, with the notable exception of recent studies performed in non-human primates[27,28].

To address these challenges, we conducted an experiment using intracranial electrical stimulation (iES) over the human cortex to determine whether iES could disrupt choice behavior and confidence judgments, using a previously validated accept/reject task[1,2]. Previous research has reported that iES of the ventral aIns elicits disgust behavior or depressed affect[29–32], whereas iES of the dorsal aIns elicits overwhelming ecstatic sensations or ingestive behavior, such as chewing and swallowing[30,33–35]. Building on this evidence, we posited that harnessing the high spatial resolution of the iES would allow us to uncover functional circuits causally linked to choice behavior.

In this work, we show that iES within the vmPFC and aIns causally influences decision-making and confidence. Our results reveal that stimulation effects on decision-making vary along a dorso-ventral axis within each region. Specifically, stimulation of the ventral aIns decreases risk-taking by increasing sensitivity to potential losses, while stimulation of the dorsal aIns and ventral vmPFC increases risk-taking by reducing sensitivity to losses. Additionally, stimulation of the aIns consistently decreases confidence levels among participants, irrespective of the stimulation location. These findings suggest the existence of functionally dissociated neural subregions and circuits within the vmPFC and aIns that are causally involved in decision-making and confidence.

Results

The effect of intracranial electrical stimulation (iES) on choice and confidence was assessed in a cohort of 15 participants (aged 34.9 ± 2.7 years, including 7 females, Supplementary Table 1) with epilepsy who had intracerebral electrodes implanted to localize their epileptogenic zone. During the choice phase of the task, participants were instructed to decide whether to accept the upcoming challenge for the stakes offered or to decline the offer and play for minimal monetary gains/losses. Additionally, participants were required to provide prospective confidence judgments regarding their task performance (Fig. 1a, b). We examined the effect of iES across a maximum number of intracerebral sites, resulting in data collected from 54 stimulation sites where iES was applied to either the aIns (Fig. 1c; $n = 38$ stimulation sites from 13 participants) or the vmPFC (Fig. 1c; $n = 16$ stimulation sites from 9 participants).

Behavioral performance

The behavioral results from the trials without iES (Fig. 1d) were consistent with the data obtained from healthy participants performing this task (Supplementary Fig. 2) or a similar version of an accept/reject choice task[2]. Specifically, the acceptance rate of participants was positively correlated with the amount of potential monetary gain and negatively correlated with the potential monetary loss ($\beta_{gain} = 0.11$,

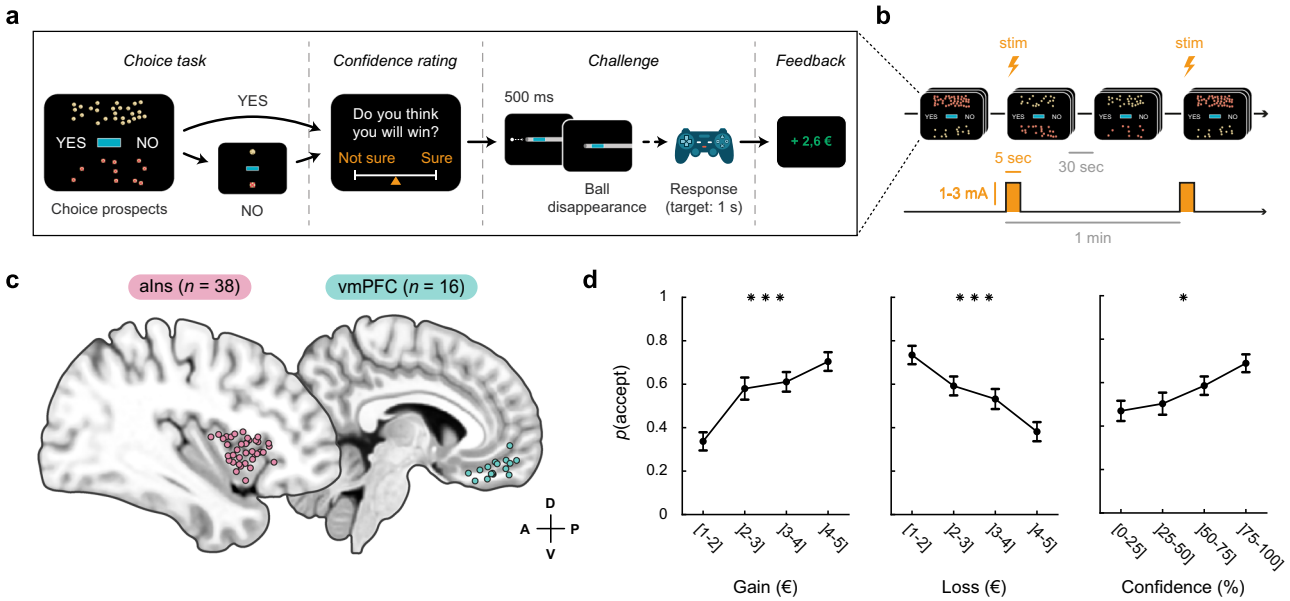

Fig. 1 | Experimental design and choice behavior. a Trial structure. Each trial consisted of a choice task and a confidence rating, followed by a challenge and feedback. In the choice task, participants had to decide whether to accept or reject a given offer based on gain prospects (represented by a set of regular 10 cent coins) and loss prospects (crossed-out 10 cent coins). The challenge involved stopping a moving ball, which became invisible upon entering a gray tunnel, within the blue target (difficulty level) in the center of the screen. **b** Stimulation procedure. Each session consisted of 14 trials with intracranial electrical stimulation (iES) alternating with 14 trials without iES. The nature of the first trial (with or without iES) was randomly assigned to maintain participants' blindness to the experimental conditions. A 30-s interval was observed between the onset of each trial, resulting in a 1 min interval between each stimulation. During trials involving iES, stimulation was delivered for a period of 5 s, in the form of pulses with a width of 0.5 ms, at a frequency of 50 Hz and an amplitude of 1, 2, or 3 mA (adjusted to ensure that participants did not experience any discernible clinical effects or subjective sensations during the task). **c** Anatomical location of the anterior insula (aIns; pink) and ventromedial prefrontal cortex (vmPFC; green) stimulation sites retained for analysis on the standard Montreal Neurological Institute (MNI) template brain. Anterior (A), posterior (P), dorsal (D) and ventral (V) directions are indicated. Stimulation sites have been aggregated in the mediolateral direction (*x*-axis) for visualization purposes. See Supplementary Fig. 1 and Supplementary Table 1 for a more detailed location of each site. **d** Choice behavior in trials without iES. Acceptance probability is plotted as a function of the monetary prospects (gain and loss) and confidence (expressed as percentile). Circles represent binned data averaged across stimulation sites. Error bars represent S.E.M across stimulation sites. Significant main effects of gains ($t_{772} = 4.24$, $p < 0.001$), losses ($t_{772} = -4.26$, $p < 0.001$), and confidence ($t_{758} = 2.23$, $p = 0.026$) on acceptance probability were found using logistic mixed-effects models with participants ($n = 15$ for gains and losses / $n = 14$ for confidence) and stimulation sites ($n = 54$ for gains and losses / $n = 53$ for confidence) as random effects. $^{*}p < 0.05$, $^{***}p < 0.001$.

https://doi.org/10.1038/s41467-024-51822-8

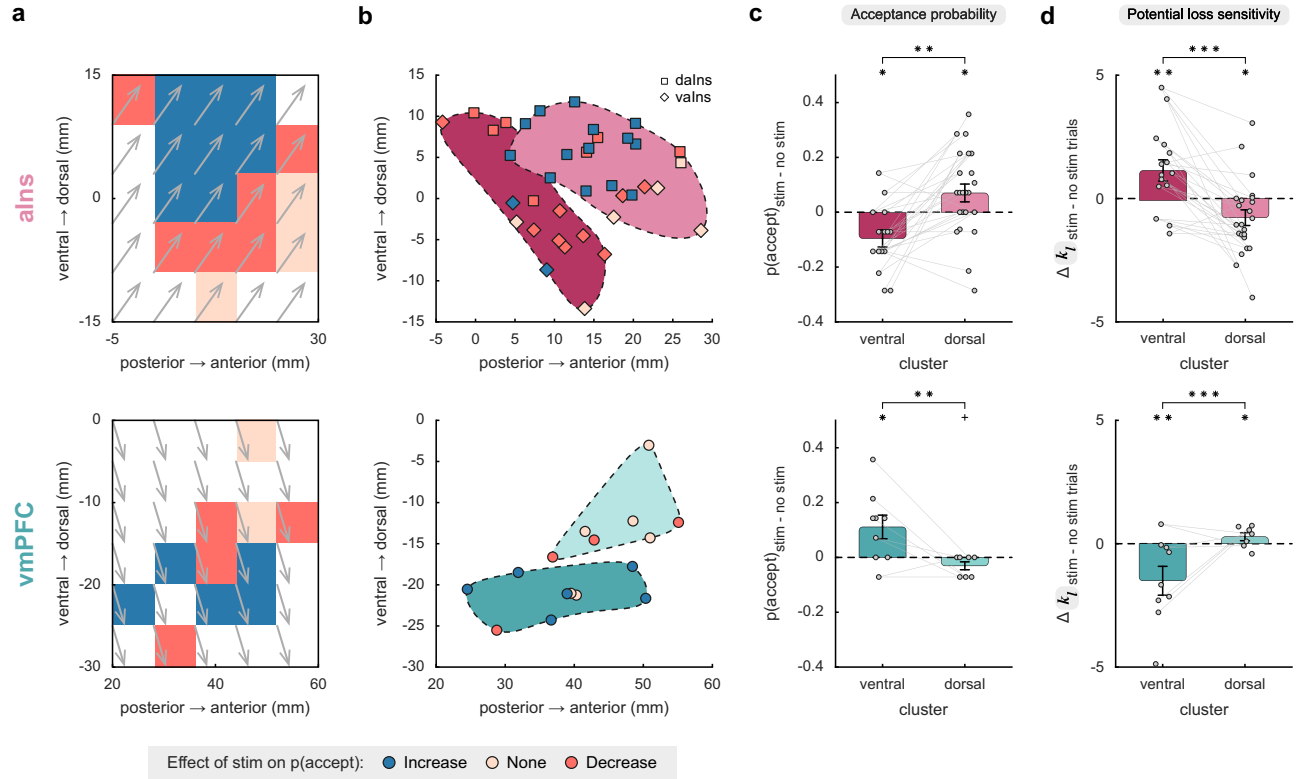

Fig. 2 | Stimulation effects on acceptance probability and computational mechanisms. a Spatial gradient of the effect of intracranial electrical stimulation (iES) on acceptance probability along the anteroposterior (*y*-axis) and ventrodorsal (*z*-axis) directions in the anterior insula (aIns; top) and ventromedial prefrontal cortex (vmPFC; bottom). Gray arrows indicate the gradient direction. **b** Clusters obtained by k-means clustering are plotted along the anteroposterior and ventrodorsal axes. In the aIns, squares represent the dorsal aIns (daIns) and diamonds represent the ventral aIns (vaIns), as delineated by Destrieux's parcellation scheme[36]. This is to compare statistically defined clusters (depicted in this figure) with anatomically defined subregions (Fig. 3). Additionally, to prevent any confusion with the anatomical terminology referring to the anterior and posterior portions of the insula, we renamed the two statistically defined clusters in this region as dorsal and ventral, respectively. **c** Effects of iES on clusters. **d** Effects of iES on participants' sensitivity to potential losses (k_l). In panels (**c**) and (**d**), dots respectively represent individual differences in acceptance probability and model weight (i.e., posterior parameter) k_l between trials with and without iES for each stimulation site. Gray lines connect the stimulation sites of the same participant. Bars and error bars represent mean and SEM, respectively, across stimulation sites. Significant interactions between stimulation condition and statistically defined clusters were found for iES effect (**c**; aIns: $F_{1,72} = 11.6$, $p = 0.001$; vmPFC: $F_{1,28} = 10.1$, $p = 0.004$) and sensitivity to potential losses k_l (**d**; aIns: $F_{1,72} = 15.3$, $p < 0.001$; vmPFC: $F_{1,28} = 13.6$, $p < 0.001$) using a linear mixed-effects model with participants (aIns: $n = 13$; vmPFC: $n = 9$) and stimulation sites (aIns: $n = 38$; vmPFC: $n = 16$) as random effects. In the aIns, post-hoc analyses reveal that iES decreased risk-taking (**c**; $F_{1,72} = 6.7$, $p = 0.012$) and increased k_l (**d**; $F_{1,72} = 7.6$, $p = 0.008$) when applied to the ventral part (dark pink; $n = 15$), while iES increased risk-taking (**c**; $F_{1,72} = 5.4$, $p = 0.023$) and decreased k_l (**d**; $F_{1,72} = 6.3$, $p = 0.014$) when applied to its dorsal part (light pink; $n = 23$). In the vmPFC, post-hoc analyses reveal that iES increased risk-taking (**c**; $F_{1,28} = 6.5$, $p = 0.016$) and decreased k_l (**d**; $F_{1,28} = 8.9$, $p = 0.006$) when applied to the ventral part (dark green; $n = 9$), whereas iES increased k_l (**d**; $F_{1,28} = 5.2$, $p = 0.030$) when applied to the dorsal part (light green; $n = 7$). Stars between bars indicate significant interactions. Stars above a single bar indicate significant post-hoc analyses (adjusted using the Benjamini-Hochberg procedure). + $p = 0.056$, *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

$t_{772} = 4.24$, $p < 0.001$, 95% CI = [0.06, 0.16]; $\beta_{loss} = -0.09$, $t_{772} = -4.26$, $p < 0.001$, 95% CI = [−0.14, −0.05]; logistic mixed-effects regression). In contrast, the challenge difficulty did not significantly modulate choice behavior ($\beta_{diff} = 7.10^{-3}$, $t_{772} = 0.33$, $p = 0.741$, 95% CI = [−0.04, 0.05]), as we experimentally set it to levels where participants' performance was close to chance. Furthermore, confidence ratings reported during trials without iES were positively associated with the probability of accepting the challenge ($\beta_{conf} = 0.39$, $t_{758} = 2.23$, $p = 0.026$, 95% CI = [0.05, 0.73]; logistic mixed-effects regression), indicating reliable metacognitive judgments.

Opponent effects of iES on choice revealed within aIns and vmPFC subregions

A preliminary analysis of behavioral data suggested that the effect of iES on behavior may vary depending on its spatial location within the aIns or vmPFC regions (Fig. 2a). To quantify this, we investigated whether the spatial coordinates of iES (in the Montreal Neurological Institute space) could predict choice. The results revealed that the location of iES had a significant impact on the choices made within both brain regions. In the aIns, a significant effect was observed along the antero-posterior (*y*) axis ($\beta_y = 0.04$, $t_{34} = 2.15$, $p = 0.039$, 95% CI = [2.10^{-3}, 0.08]), as well as a trend along the ventro-dorsal (*z*) axis ($\beta_z = 0.04$, $t_{34} = 1.86$, $p = 0.072$, 95% CI = [4.10^{-3}, 0.08]). Similarly, in the vmPFC, the ventro-dorsal (*z*) localization of iES had a significant influence on choices made ($\beta_z = -0.10$, $t_{12} = -3.71$, $p = 0.003$, 95% CI = [−0.16, −0.04]). These findings were replicated using K-means clustering on the data (Fig. 2b), which demonstrated the existence of two distinct areas within each brain region. In the aIns, this was a postero-ventral cluster with $n = 15$ sites and an antero-dorsal cluster with $n = 23$ sites. In the vmPFC, this was a ventral cluster with $n = 9$ sites and a dorsal cluster with $n = 7$ sites.

In the following section, we further specify the impact of iES on economic choices (Fig. 2c). In the aIns, we observed a significant interaction between stimulation condition and statistically defined clusters ($\beta = -0.041$, $F_{1,72} = 11.6$, $p = 0.001$, 95% CI = [−0.06, −0.02]; linear mixed-effects model). Post-hoc analyses revealed that iES decreased risk-taking when applied to the ventral part of the aIns ($F_{1,72} = 6.7$, $p = 0.012$), whereas iES increased risk-taking when applied

to the dorsal part of the aIns ($F_{1,72} = 5.4$, $p = 0.023$). Similarly, we found a significant interaction between stimulation condition and vmPFC clusters ($\beta = -0.038$, $F_{1,28} = 10.1$, $p = 0.004$, 95% CI = [−0.06, −0.01]). Post-hoc analyses revealed a pattern opposite to the aIns findings, with iES increasing the probability of accepting the challenge when applied to the ventral part of the vmPFC ($F_{1,28} = 6.5$, $p = 0.016$), while it tended to decrease risk-taking when applied to the dorsal part of the vmPFC ($F_{1,28} = 4.0$, $p = 0.056$). Furthermore, we conducted control analyses to check the robustness of the effects of iES on choice across participants, by performing a participant ablation analysis (Supplementary Fig. 3), in which all stimulation sites of a participant were iteratively removed from the dataset (see also Supplementary Fig. 4 for individual participant anatomical and choice data). Finally, we tested more complex statistical models, including the magnitude of gains and losses, which suggested that the opponent effects of iES on choice in the dorsal versus ventral clusters were not influenced by the magnitude of gains or losses, either in the aIns or in the vmPFC (Supplementary Fig. 5).

iES modulated participants' subjective sensitivity to the prospect of monetary loss

The next step was to investigate the computational mechanisms through which iES disrupted choice behavior. To do this, we fitted the choice data with a model based on expected utility theory[1,2]. Our findings revealed that iES modulated participants' sensitivity to potential losses (captured by the parameter k_l) in both brain regions (Fig. 2d), but not their sensitivity to potential gains (Supplementary Fig. 6). Specifically, we observed a significant interaction between stimulation condition and clusters regarding the effect on k_l in both the aIns ($\beta = 0.536$, $F_{1,72} = 15.3$, $p < 0.001$, 95% CI = [0.26, 0.81]) and vmPFC ($\beta = 0.541$, $F_{1,28} = 13.6$, $p < 0.001$, 95% CI = [0.24, 0.84]). Subsequent post-hoc analyses revealed that iES applied to the ventral part of the aIns increased k_l ($F_{1,72} = 7.6$, $p = 0.008$), while iES applied to its dorsal part decreased k_l ($F_{1,72} = 6.3$, $p = 0.014$). Again, an opposite dorso-ventral gradient was identified within the vmPFC, whereby iES in its ventral part decreased k_l ($F_{1,28} = 8.9$, $p = 0.006$), while iES in its dorsal part increased participants' sensitivity to monetary loss prospects ($F_{1,28} = 5.2$, $p = 0.030$).

Anatomical specification of functional parcellation by iES of aIns and vmPFC

To exclude the possibility that the previous results were partly biased by a double-dipping issue, stemming from the use of K-means clustering to delineate functional subregions within either the aIns or vmPFC, we proceeded to define dorsal or ventral functional subregions within the aIns or vmPFC using individual anatomical landmarks (Fig. 3a). For the anterior insula, we used Destrieux's atlas[36] (Supplementary Fig. 7) to differentiate between iES sites located within the short insular gyri and the region of the insula bounded by the anterior circular sulcus (corresponding to the ventral aIns; vaIns) and those located within the region of the insula bounded by the superior circular sulcus (corresponding to the dorsal aIns; daIns). iES applied to these anatomically defined subregions of the aIns differentially affected participants' choices in the same directions as in the previous analysis (Fig. 3b top; interaction: $\beta = 0.04$, $F_{1,72} = 12.4$, $p < 0.001$, 95% CI = [0.02, 0.07]). Specifically, iES increased risky decisions in the dorsal aIns (post-hoc: $F_{1,72} = 6.58$, $p = 0.012$), whereas iES decreased risk-taking in the ventral aIns (post-hoc: $F_{1,72} = 6.59$, $p = 0.012$). Similar results were found for participants' sensitivity to loss (Fig. 3c top; interaction: $\beta = -0.47$, $F_{1,72} = 11.9$, $p = 0.001$, 95% CI = [−0.74, −0.20]), where iES in the dorsal aIns decreased k_l ($F_{1,72} = 5.3$, $p = 0.025$), while iES in the ventral aIns increased k_l ($F_{1,72} = 6.5$, $p = 0.013$). Interestingly, the opponent effects of iES on choice along the dorso-ventral axis in the aIns were successfully replicated using an alternative parcellation scheme[37] (the Julich Brain Atlas), thereby revealing a significant interaction between stimulation condition and Julich's defined daIns and vaIns subregions ($\beta = 0.038$, $F_{1,64} = 4.6$, $p = 0.036$, 95% CI = [2.10^{-3}, 0.07]; linear mixed-effects model; Supplementary Fig. 8). Unfortunately, the spatial sampling from the anatomically defined dorsal part of the vmPFC (namely the suprarostral sulcus; SU-ROS) was not sufficient ($n = 3$) to perform a similar analysis for the vmPFC. Nevertheless, upon evaluating the effect of iES solely on the ventral part of the vmPFC (recordings within the superior rostral sulcus; ROS-S), we observed that iES increased risky-choices (Fig. 3b bottom; $\beta = -0.039$, $F_{1,20} = 4.5$, $p = 0.046$, 95% CI = [−0.08, -8.10^{-3}]; linear mixed-effects model) and decreased participants' sensitivity to potential losses (k_l; $\beta = 0.59$, $F_{1,20} = 5.0$, $p = 0.037$, 95% CI = [0.04, 1.14]; Fig. 3c bottom), consistent with results based on statistically defined vmPFC subregions (K-means clustering approach, see Fig. 2).

iES of aIns reduced prospective confidence judgements

Finally, we examined the influence of iES on confidence levels (Fig. 4). In the aIns, a significant negative main effect of iES on confidence scores was observed ($\beta = 0.06$, $F_{1,70} = 10.3$, $p = 0.002$, 95% CI = [0.02, 0.11]; linear mixed-effects model), with no dissociable effect of iES on confidence in its dorsal and ventral parts (Supplementary Fig. 9; non-significant interaction effect $\beta = 0.01$, $F_{1,70} = 0.14$, $p = 0.71$, 95% CI = [−0.04, 0.07]). This indicates that applying iES over the entire aIns significantly reduced participants' confidence compared to trials without iES. This effect was not explained by a possible impact of iES on performing the challenge ($\beta = -4.10^{-3}$, $F_{1,72} = 0.12$, $p = 0.73$, 95% CI = [−0.03, 0.02]). In the vmPFC, no significant interaction ($\beta = -0.04$, $F_{1,28} = 0.51$, $p = 0.48$, 95% CI = [−0.14, 0.07]) or main effect of stimulation ($\beta = -4.10^{-3}$, $F_{1,28} = 0.01$, $p = 0.90$, 95% CI = [−0.07, 0.08]) and cluster ($\beta = 2.10^{-11}$, $F_{1,28} = 4.10^{-11}$, $p = 1$, 95% CI = [-6.10^{-6}, -6.10^{-6}]) was found, suggesting that iES on this brain region did not affect participants' confidence levels.

Discussion

An emerging consensus suggests that the vmPFC and aIns may play opposing roles in economic decisions, mood, and learning processes. Specifically, neural activity in the vmPFC is associated with increased risk-seeking, pleasantness, reward-based learning, and good mood[1,2,38–42]. In contrast, neural activity in the aIns has been linked to risk-averse decisions, unpleasantness, punishment avoidance, and bad mood[1,2,39,41,43–47]. Our findings provide evidence for the causal involvement of functional subregions within the aIns and vmPFC in decision-making and confidence processes. We found that stimulation of the ventral aIns reduced risk-taking by increasing participants' sensitivity to monetary loss prospects, whereas stimulation of the ventral vmPFC and dorsal aIns increased risk-taking by decreasing participants' sensitivity to potential loss. Additionally, iES applied to the aIns decreased confidence levels, whereas similar stimulation in the vmPFC did not disrupt confidence judgments.

In the aIns, we have demonstrated that applying iES to the dorsal portion of the anterior insula, bounded by the superior circular sulcus, led to an underestimation of loss prospects, thereby increasing risk-taking behavior. Conversely, stimulation of the ventral aIns (specifically, the short insular gyri and the region bounded by the anterior circular sulcus) led to an overestimation of loss prospects, resulting in a decrease in risk-taking behavior. These findings in humans align with previous research in mice, indicating distinct areas within the insula governing approach and avoidance behaviors[48,49]. Our results also corroborate findings from stimulation studies in non-human primates, where stimulation of the ventral aIns reduced approach behaviors in appetitive contexts[50] and was associated with negative feelings such as disgust[29–31] or depression[32]. In contrast, iES of the dorsal aIns has been more commonly associated with positive behaviors such as ingestive behaviors in monkeys[30] or even ecstatic feelings in rare human case studies[33–35]. Therefore, by combining iES with an economic choice paradigm, our study bridges the gap between previous human and

https://doi.org/10.1038/s41467-024-51822-8

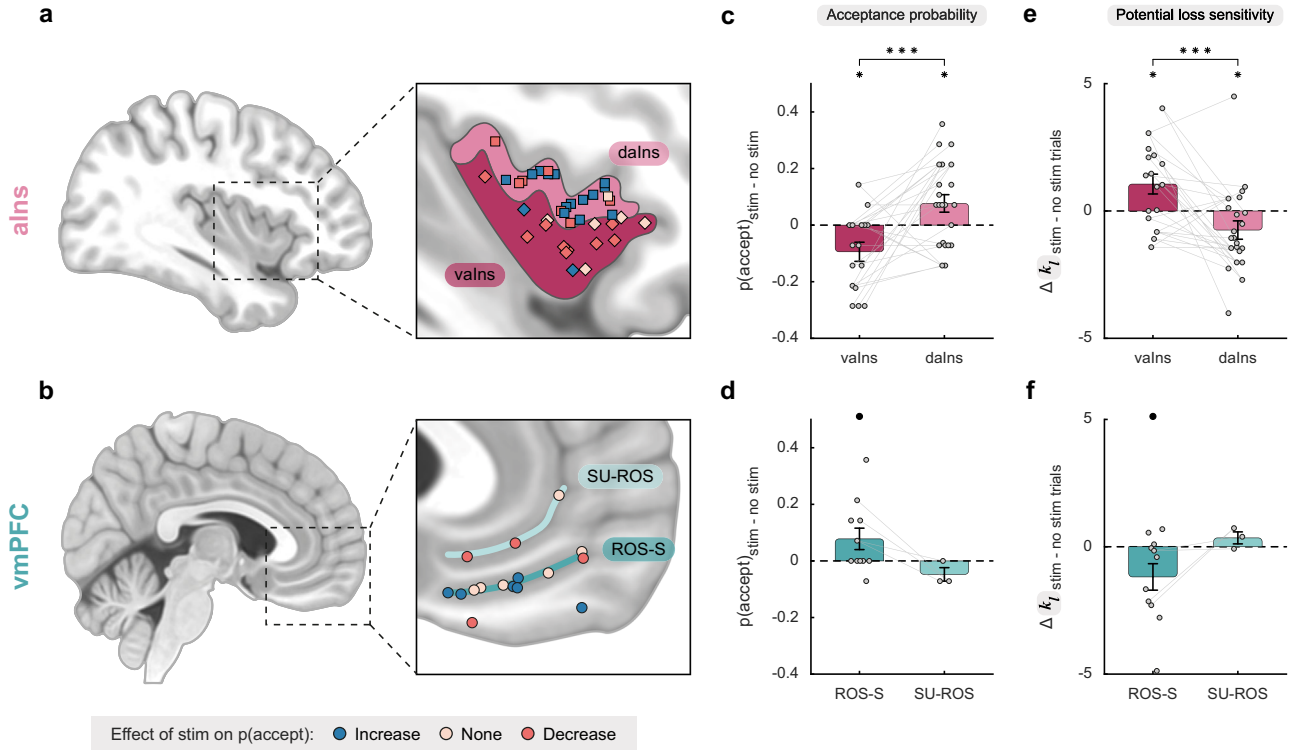

Fig. 3 | Effects of stimulation on acceptance probability and computational mechanisms in anatomically defined subregions. a Location of stimulation sites in the anterior insula (aIns) according to Destrieux's parcellation scheme[36] (Supplementary Fig. 7). The ventral anterior insula (vaIns; diamond) is defined as parcels 18 and 47 of Destrieux's atlas, corresponding to the short insular gyri and the region of the insula bounded by the anterior circular sulcus, respectively. The dorsal anterior insula (daIns; square) is defined as parcel 49 of Destrieux's atlas, corresponding to the part of the insula that is bounded by the superior circular sulcus. This partitioning based on individual anatomy resulted in groupings very close to the previous clusters (over 76% similarity between the two methods). **b** Location of stimulation sites according to the two main sulci of the ventromedial prefrontal cortex (vmPFC): the suprarostral sulcus (SU-ROS) and the superior rostral sulcus (ROS-S)[71]. Each electrode was manually positioned on an MNI template based on individual anatomy (see Supplementary Fig. 1 for the location of each site on individual anatomies). **c**–**d** Effects of iES on acceptance probability in aIns (**c**) and vmPFC (**d**). **e**–**f** Effects of iES on participants' sensitivity to potential losses (k_l) in aIns (**e**) and vmPFC (**f**). In panels (**c**) to (**f**), dots represent the individual differences in acceptance probability or model weight (i.e., posterior parameter) k_l between trials with and without iES for each stimulation site. Gray lines connect the stimulation sites of the same participant. Bars and error bars represent mean and SEM, respectively, across stimulation sites. In the aIns, significant interactions between stimulation condition and anatomical subregions were found for iES effect (**c**; $F_{1,72} = 12.4$, $p < 0.001$) and sensitivity to potential losses k_l (**e**; $F_{1,72} = 11.9$, $p = 0.001$) using a linear mixed-effects model with participants ($n = 13$) and stimulation sites ($n = 38$) as random effects. Post-hoc analyses show that iES decreased risk-taking (**c**; $F_{1,72} = 6.59$, $p = 0.012$) and increased k_l (**e**; $F_{1,72} = 6.5$, $p = 0.013$) in the vaIns, while iES increased risk-taking (**c**; $F_{1,72} = 6.58$, $p = 0.012$) and decreased k_l (**e**; $F_{1,72} = 5.3$, $p = 0.025$) in the daIns. In the ROS-S region of the vmPFC, significant main effects of stimulation condition were found for iES effect (**d**; $F_{1,20} = 4.5$, $p = 0.046$) and sensitivity to potential losses k_l (**f**; $F_{1,20} = 5.0$, $p = 0.037$) using a linear mixed-effects model with participants ($n = 8$) and stimulation sites ($n = 11$) as random effects. Stars between bars indicate significant interactions. Stars above a single bar indicate significant post-hoc analyses (adjusted using the Benjamini-Hochberg procedure). Black dots above a single bar indicate a significant main effect of stimulation condition on the corresponding subregion (model without interaction term). $^{*}p < 0.05$, $^{***}p < 0.001$.

preclinical evidence by demonstrating dissociable effects of iES within the aIns in a single cohort of participants.

Our findings indicate that iES in the aIns is associated with a decrease in confidence, regardless of its specific localization within this region. This provides causal evidence for the critical involvement of the aIns in metacognitive judgments. These results extend those of previous studies that demonstrated a negative correlation between neural activity in the aIns and confidence[7–10,51]. Relatedly, recent evidence indicates that value and confidence interact at both behavioral and neural levels[11,51,52]. Notably, higher monetary gains have been shown to bias confidence judgments upwards, whereas higher monetary losses bias confidence downward. Accordingly, our findings regarding the vaIns are consistent with this view, as iES in the aIns decreased both participants' sensitivity to losses and their confidence levels. Conversely, the antagonist effect of iES on choices and confidence in the daIns challenges this view and was more surprising, suggesting that neural processes involving the daIns and associated brain networks might implement confidence and choice processes independently.

In the vmPFC, our results indicate that stimulation of the superior rostral sulcus (ROS-S) led to an underestimation of loss prospects, resulting in an increase in risk-taking. Since existing studies investigating the effects of stimulation of this area have often reported null results[53–55], we propose that combining iES with cognitive paradigms designed to tease apart precise decision-making functions paves the way for future research aimed at delineating the causal role of the vmPFC more precisely. However, while we were able to replicate the dorso-ventral functional dissociation in the aIns using either statistical or anatomically defined vaIns and daIns, the results were less reliable for the vmPFC. Although our results also suggest the existence of subregions within the vmPFC, we lacked sufficient statistical power to investigate the effects of iES on the anatomically defined dorsal part of the vmPFC (i.e. the suprarostral sulcus, SU-ROS; 3 sites; Supplementary Table 1) versus its more ventral sulci (11 sites in ROS-S and 2 sites in the most ventral part of the vmPFC; Supplementary Table 1). Consequently, future work is needed to further specify the functional parcellation between the dorsal and ventral subregions of the vmPFC.

https://doi.org/10.1038/s41467-024-51822-8

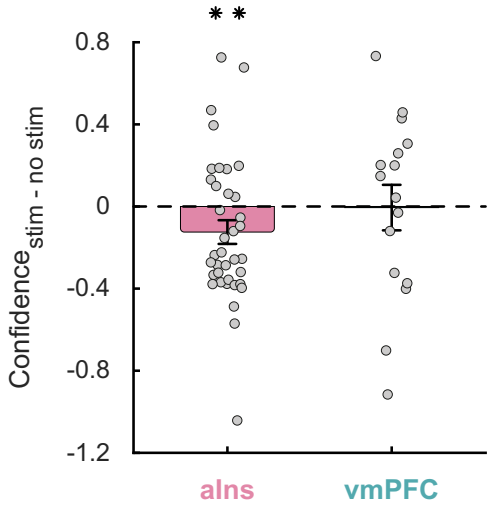

Fig. 4 | Stimulation effects on confidence. Dots represent individual difference in confidence ratings between trials with and without intracranial electrical stimulation for each stimulation site in the anterior insula (aIns; left) and ventromedial prefrontal cortex (vmPFC; right). Bars and error bars respectively represent mean and SEM across stimulation sites. The stars indicate the main effect of stimulation ($F_{1,70} = 10.3$, $p = 0.002$), obtained when examining the effect of stimulation conditions and clusters on confidence ratings using a linear mixed-effects model with participants ($n = 12$) and stimulation sites ($n = 37$) as random effects. $^{**}p < 0.01$.

Another unexpected finding was that, despite the existing evidence positively associating the neural activity in the vmPFC with confidence judgments across various tasks[3–5,51,56–58], we were unable to demonstrate a causal role of the vmPFC for confidence. Given these previous findings, we initially expected that iES applied to the vmPFC would increase confidence. Although negative findings should be interpreted with caution, we can only speculate that metacognitive judgments may involve a higher level of variability or noise. This, in conjunction with the fact that confidence ratings were temporally delayed to the iES onset compared to choices, may have impeded the emergence of a statistically significant effect of iES on the vmPFC regarding confidence.

Interestingly, we observed similar effects of iES on choice behavior and participants' sensitivity to loss prospects for both the vmPFC and daIns. However, it is still possible that these effects may be mediated by distinct cognitive processes. For instance, given that neural activity in the vmPFC also increases during self-referential episodic memory retrieval[59], the increase in risky choices observed after iES on the vmPFC in our study could be related to a reduced sensitivity to emotional markers of self-referential risk. Additionally, previous studies have shown that intracranial activity in the aIns is modulated by negative outcomes or decision variables associated with monetary losses[1,39,60], whereas intracranial activity in the vmPFC is associated with monetary gains or positive outcomes[1,39].

In a previous study employing a similar choice task while recording intracranial EEG activity[1], we demonstrated that increased broadband gamma activity (BGA, 50-150 Hz) in the vmPFC or vaIns promoted risk-taking behavior, while increased BGA in the daIns tempered risky decision-making. These electrophysiological findings align with the causal findings of the present study, in that the ventral and dorsal parts of the anterior insula seem to implement dissociable roles during decision-making. However, it remains unclear why neural activity in the ventral part of the insula promotes risk-taking[1], whereas in the present study, iES increased risk-taking when it was applied to the dorsal part of the insula. Similarly, while vmPFC activity has been correlated with increased sensitivity to monetary gains[1,2], our study associated the effects of iES on the vmPFC with altered sensitivity to monetary losses. The discrepancies observed between invasive electrophysiology and stimulation studies may be partly attributable to differences in statistical power that arise from the relatively limited number of choices made by the participants. In particular, this small number of trials may have impacted our computational results since it imposed more constraints on parameter fitting compared to our previous electrophysiology study[1]. Furthermore, the effects of iES on choices could be related to the modulation of functional connectivity, given the existence of heterogeneous connectivity patterns in the dorsal versus ventral subregions of aIns[61–63] and vmPFC[64–66]. However, the limited number of trials, low spatial sampling within the same patient and electrical artefacts induced by stimulation precluded any meaningful investigation of the local and network-based electrophysiological mechanisms underlying the effects of stimulation on our regions of interest.

Of course, iES data can only be derived from individuals with severe drug-resistant epilepsy, yet we interpreted our data as if they were from healthy participants for several reasons. Firstly, the participants' choice behavior mirrored that of healthy individuals in similar tasks[2]. Additionally, the neural responses in the vmPFC or aIns observed during intracranial recordings in value-based decision-making tasks[1,5,39] coincided with brain regions identified during functional imaging in healthy participants performing the same tasks[2,45,67]. Finally, it has been demonstrated that epileptic tissue can generate normal neuronal activity during cognitive tasks[68]. Furthermore, the advantage of iES lies in its precise targeting of specific brain regions, allowing for the differentiation between the ventral and dorsal parts of the aIns or vmPFC. Achieving such sub-regional specificity would be challenging with other causal techniques available in human cognitive neuroscience.

To conclude, our study represents a step forward in elucidating the causal contribution of opponent circuits involving the vmPFC and aIns, demonstrating the existence of dorso-ventral subregions in value-based decision-making processes. Given the involvement of these circuits in several neuropsychiatric disorders, our findings offer valuable insights that can inform the ongoing development of closed-loop brain stimulation approaches, aiming to better alleviate symptoms that are otherwise resistant to existing treatments.

Methods

Patient selection

Participants were fifteen patients suffering from pharmaco-resistant focal epilepsy and candidates to surgical treatment (34.9 ± 2.7 years old, 7 females, 5 left-handed, see demographic details in Supplementary Table 1). These patients underwent a monitoring of their neural activity at the Epilepsy Monitoring Unit of Grenoble University Hospital (Grenoble, France) using stereotaxic implanted multilead deep electrodes (sEEG) as part of their pre-surgical evaluation meant to localize the epileptogenic zone that could not be identified through noninvasive methods. Electrode implantation was performed according to routine clinical procedures with all targeted structures selected strictly according to clinical considerations for the pre-surgical evaluation with no reference to the current study. Patients were included in the study if they had electrodes implanted in at least one brain regions of interest (i.e., aIns and/or vmPFC), and if they were willing and able to perform the choice task. No statistical method was used to predetermine sample size. All patients were taking anti-seizure medications (see Supplementary Table 1), some of which were reduced or stopped before stimulation sessions on clinical grounds. Exclusion criteria were age under 18 and complete inability to speak French (to prevent improper execution of the behavioral task due to misunderstanding of task instructions).

Data from nineteen healthy participants performing the same behavioral task as patients with epilepsy were also obtained for comparative analysis (Supplementary Fig. 2).

Ethical approval
All patients and healthy participants gave their informed consent to participate in the study. The study involving patients was approved by the Ethics Committee "Comité de protection des personnes Sud Ouest et Outre Mer IV" (IdRCB: 2017-A03248-45). The study involving healthy participants was approved by the local ethics committee (CERGA-Avis-2021-7).

Electrodes implantation and location
Depth electrodes were implanted using robot-assisted sEEG electrode implantation technique (ROSA robot). Fifteen to eighteen semi-rigid electrodes were implanted per patient. Each electrode had a diameter of 0.8 mm and, depending on the target structure, contained 10-18 contact leads of 2 mm wide and 1.5 mm apart (Dixi Medical, Besançon, France). Stimulation sites were chosen in consultation with neurologists for their location in one of our two regions of interest (aIns or vmPFC). To confirm their exact location, the electrodes were also anatomically labeled by co-registering a pre-operative anatomical magnetic resonance imaging (MRI, 3D T1 contrast) with a post-operative computed tomography (CT) scan obtained for each patient, using the IntrAnat Electrodes software[69]. The subregions of the anterior insula (aIns) were delineated using the Destrieux atlas[36] (Supplementary Fig. 7). The ventral anterior insula (vaIns) is defined as the short insular gyri (parcel 18; G_insular_short) and the region of the insula bounded by the anterior circular sulcus, which vertically separates it from the orbital gyri (parcel 47; S_circular_insula_ant). The dorsal anterior insula (daIns) is defined as the part of the insula that is bounded by the superior circular sulcus, which horizontally separates it from the subcentral and inferior frontal gyri (parcel 49; S_circular_insula_sup). It is important to note that parcel 49, which is designated as the daIns, also includes a portion of the posterior insula. However, since our stimulation sites were specifically chosen anterior to the central sulcus to target the anterior insula, all sites fell within the anterior portion of this region ($y > -0.2$ in MNI space). The stimulation sites in the vmPFC matched those of the same name in the MarsAtlas parcellation scheme[70], corresponding to the most medial part of the ventral prefrontal cortex (Supplementary Fig. 10). Note that this parcellation scheme distinguishes the vmPFC from the orbitofrontal cortex (OFC), as it divides the ventral prefrontal cortex into four subregions, which are (from the most medial to the most lateral): the ventromedial prefrontal cortex (vmPFC), the ventromedial orbitofrontal cortex (vmOFC), the ventral orbitofrontal cortex (vOFC), and the ventrolateral orbitofrontal cortex (vlOFC). Identification of vmPFC sulci on individual anatomies was done manually based on the literature standard[71].

Intracranial electrical stimulation
Intracranial electrical stimulations (iES) were applied between two contiguous contacts located in a region of interest. Bipolar stimuli were delivered on a pair of contacts (defined as a stimulation site) using a constant current rectangular pulse generator designed for a safe diagnostic stimulation of the human brain (Micromed, Treviso, Italy), according to parameters used in clinical procedures and proven to produce no structural damage. High-frequency stimulation at 50 Hz, with a pulse width of 0.5 ms and an intensity of 1, 2, or 3 mA, was applied in a bipolar fashion during a 5 s period on a stimulation site. For a given stimulation site, the stimulation intensity was determined as the highest intensity devoid of clinical symptoms or electroclinical responses during previous clinical stimulation sessions as systematically done in our routine procedure. To clarify, while the maximum intensity of iES used at Grenoble University Hospital is 3 mA, we deliberately selected an intensity of 2 mA for 8/54 sites, and 1 mA for 7/54 sites because clinical effects were observed at 3 mA and 2 mA, respectively, for these stimulation sites during routine clinical stimulations (see Supplementary Table 2).

Stimulation sessions
Each session of the experiment corresponded to the stimulation of a specified stimulation site at a given intensity. Thus, the number of sessions performed by each patient was determined by the number of contact pairs available in regions of interest. Each experimental session consisted of 28 trials of a behavioral task alternating between stimulation ($n = 14$) and non-stimulation ($n = 14$) trials (Fig. 1b). The first trial of the session was randomly assigned to either stimulation or non-stimulation to maintain patients' blindness on experimental conditions. Stimulations were triggered manually and iEEG activity was monitored in real-time to detect stimulation-induced after discharges and electrographic seizures, and to stop the session immediately if necessary. The between trials time interval was also kept above 30 s, so that two iES remained separated by a minimum interval of 1 min. In stimulation trials, the stimulation was initiated about 1 s before the trial onset.

Behavioral task
Presentation of visual stimuli and acquisition of behavioral data were performed on a PC using custom Matlab scripts implementing the PsychToolBox libraries[72]. All participant responses were done with a gamepad (Logitech F310S) using both hands. Before the experiment, patients were informed that various brain regions would be stimulated, but they were blind to experimental conditions, including stimulation parameters, the brain region being tested, and whether stimulation was active. Each trial consisted of a choice task combined with confidence rating and a challenge (Fig. 1).

Choice task. The choice task began with the presentation of an offer consisting of three attributes: a gain prospect (represented by a bunch of 10-cent coins, range: 1–5 €), a loss prospect (represented by crossed out 10 cent coins, range: 1–5 €) and the upcoming challenge difficulty (represented by a target window corresponding to ~50% theoretical success). Challenge difficulty was displayed on the center of the screen (see training section for further details about how difficulty was adjusted to each participant). Participants were asked to accept or reject this offer by pressing a left or right button depending on where the choice option ("yes" or "no") was displayed. Participants' choice determined the amount of money at stake: accepting meant that they would eventually win the gain prospect or lose the loss prospect based on their performance in the upcoming challenge, whereas declining the offer meant playing the challenge for minimal stakes (winning 10 cents or losing 10 cents). The sequence of trials was pseudo-randomized such that all delta values, computed as the difference between gain and loss prospects, continuously sampled along ten intervals ([−40 −30], [−30 −20], [−20 −10], [−10 −5], [−5 0], [0 5], [5 10], [10 20], [20 30], [30 40]), were displayed for one patient during a session, with the four medium intervals ([−10 −5], [−5 0], [0 5] and [5 10]) presented twice ($n = 8$ out of 14 trials) to maximize the occurrence of difficult choices. The positions of gain and loss prospects were randomly determined to be either displayed on top or bottom of the screen and similarly, the choice options ("yes" or "no") were randomly displayed on the left or right. Stimulation and non-stimulation trials were strictly identical in one session, such that participants served as their own control. Patients had a free time delay to accept or decline the offer. If they declined the offer, a 250 ms screen displayed the new offer (a minimal stake of 10 cents). Thus, the challenge was performed regardless of the choice answer to prevent patients from eventually rejecting more offers to decrease experiment duration.

Confidence ratings. Before performing the challenge, patients were asked to rate their confidence in winning the challenge by answering the following question "Do you think you will win?". Patients had a free time delay to answer by moving a cursor from left (not sure) to right (sure) along a continuous visual analog scale (100 steps) with left- and

right-hand response buttons. The initial position of the cursor on the scale was randomized to avoid confounding confidence level with movements' quantity.

Challenge. The challenge started right after confidence confirmation: a ball appeared on the left of the screen and moved, horizontally and at constant speed, towards screen center. Patients were asked to press the confirmation button when they thought the ball was inside the box displayed at screen center (i.e., the target window). The ball always reached the center of the target after 1 s. Thus, the size of the target window represented the margin of error tolerated in patient reaction time (target: 1 s after the movement onset of the ball). Unbeknownst to the patients, the success rate was maintained at about 50% by decreasing (if the challenge was successful) or increasing (if the challenge was unsuccessful) the tolerated margin of error by 1% theoretical success after each trial. Importantly, the moving ball could only be seen during the first 500 ms (half of the trajectory), and patients had to extrapolate the last 500 ms portion of the ball's trajectory to assess whether the ball was inside the target. Finally, feedback of 1 s was given to the patients about their payoff after the challenge. Also note that, to improve patients' motivation to perform the task as accurately as possible, the total amount of money earned by the patients during a session (calculated by adding gains and losses across all trials) was displayed at the end of a session.

Training. Before the main experiment, a training—divided into three steps—familiarized participants with all sub-parts of the task. In the first step, they were familiarized with the challenge by performing 30–80 trials of it. Unbeknownst to the participants, the size of the bar was decreased after each success and increased after each error by one pixel, so that their performance statistically converged at 50% success. This first task was completed when participants' performance stabilized (most often before <80 trials). Each training trial was followed by feedback informing whether the challenge was successful ("ok" in green) or missed ("too slow" or "too fast" in red). In the second step, participants completed 64 trials of the full choice (i.e., the challenge was always preceded by an offer), and feedback on the money won/lost in the trial was displayed at the end of each trial. The goal was to train patients to properly integrate the dimensions of the offer (gains and losses) when making their choice. Finally, the third and last part of the training (8 trials) was completely similar to the main task to allow participants to familiarize with confidence ratings.

Another purpose of the training was to tailor the difficulty of the challenge to each participant's abilities. To do this, we calculated the mean and standard deviation of challenge performance across all training trials (including all three steps) for each participant, assuming that errors were normally distributed. We then computed an individual tolerated margin of error, corresponding to a theoretical success of 55%, to be used on the first trial of a session of the main task. Among patients, the tolerated margins of error corresponding to 55% theoretical success ranged from [± 40 ms] in the most precise patient to [± 206 ms] in the less precise one.

Behavioral analysis

Statistical analyses were performed with Matlab Statistical Toolbox (Matlab R2018b, The MathWorks, Inc., USA), and generalized linear mixed-effect (glme) models were estimated using the "fitglme" function with default parameters that maximize the maximum pseudo-likelihood of the observed data under the model. Assumptions of each glme, including linearity, normality, homoscedasticity, absence of autocorrelation, and absence of multicollinearity, were thoroughly assessed prior to further analyses. Post-hoc tests were performed using the Matlab function "coefTest". The p-values for these comparisons were then adjusted using the Benjamini-Hochberg procedure to control the false discovery rate (FDR) at $\alpha = 0.05$.

Experimental sessions in which participants systematically accepted or declined the offer on all trials, irrespective of the stimulation condition, were excluded from the analyses. Consequently, out of the 59 sessions recorded in 15 participants, 54 sessions were retained for analysis (Supplementary Table 1).

Choice behavior. Analysis of choice behavior was performed on non-stimulated trials across all sessions (i.e., stimulation sites) using glme models that included choice as the dependent variable (modeled using a binomial response function distribution, i.e., logistic regression) and (1) challenge difficulty, gain and loss magnitude, or (2) confidence as predictor variables. The models also comprised a full random-effects structure at the participant and stimulation site levels (i.e., intercepts and slopes for all predictor variables). For example, in Wilkinson-Rogers notation, the glme model of the effect of offer dimensions on choices can be written as follows:

$$\begin{aligned} choice \sim 1 + gain + loss + diff + (1 + gain + loss + diff \,|\, participant) \\ + (1 + gain + loss + diff \,|\, stimulation\ site) \end{aligned} \quad (1)$$

For each fixed effect, the estimated coefficient value (β) is reported, as well as the t-statistic (testing that the coefficient is equal to 0) along with its p-value and 95% confidence interval (CI).

Spatial gradient of the effect of stimulation on acceptance probability. This glme model included the effect of stimulation on acceptance probability as the dependent variable (coded as −1, 0, or 1), the MNI coordinates x, y, and z as predictor variables, and a full random-effects structure at the participant level (i.e., intercepts and slopes for all predictor variables). For each fixed effect, the estimated coefficient value (β) is reported, as well as the t-statistics (testing that the coefficient is equal to 0) along with its p-value and 95% CI.

Interaction effects of stimulation and clusters/subregions on choice/confidence. For each region of interest (ROI), these glme models included (1) acceptance probability across trials of the same stimulation condition, or (2) confidence ratings as the dependent variable. The models also comprised main effects and interaction of the stimulation condition and clusters (or subregions) as predictor variables, a full random-effects structure at the participant level (i.e., intercepts and slopes for all predictor variables), and the main effect of stimulation condition at the stimulation sites level. For example, in Wilkinson-Rogers notation, the glme model of interaction effects between stimulation conditions and clusters on choice can be written as follows:

$$\begin{aligned} p_{accept} \sim 1 + stim\ condition{*}cluster\ ID \\ + (1 + stim\ condition{*}cluster\ ID | participant) \\ + (1 + stim\ condition | stimulation\ site) \end{aligned} \quad (2)$$

The coding of the dummy variables was specified as "effects", so that the sum of the dummy variable coefficients is equal to 0[73]. The F-values for interactions, main effects, and post-hoc tests are reported, along with their degrees of freedom and corresponding p-values. Additionally, estimated coefficient values (β) and 95% confidence intervals (CI) are provided for interactions and main effects. Please note that one participant did not perform confidence ratings, which resulted in the use of one less stimulation site in the anterior insula to investigate the effects of stimulation on confidence (37 sessions instead of 38).

Because spatial sampling from the anatomically defined dorsal part of the vmPFC (i.e., the suprarostral sulcus; SU-ROS) was not sufficient, we also tested the effect of iES only on the ventral part of the

vmPFC (i.e., the superior rostral sulcus; ROS-S) by removing the subregion predictor from the model, yielding for example:

$$p_{accept} \sim 1 + stim\ condition + (1 + stim\ condition|participant) + (1 + stim\ condition|stimulation\ site) \quad (3)$$

Computational analysis

Choice model. Choices were fitted using a published computational framework[1,2] on additional behavioral data acquired separately from 10 of 15 participants (128 or 192 trials of the task without stimulation). Acceptance probability was calculated as a sigmoid function (softmax) of expected utility:

$$p(accept,t) = \frac{1}{1 + e^{-(utility + k_t \times t)}} \quad (4)$$

where k_t is a free parameter that accounts for a linear drift with time (trial index t) to capture fatigue effects. The utility function is based on expected utility theory where potential gains and losses are multiplied by probability of success (p_s) vs. failure ($1 - p_s$):

$$utility = k_0 + p_s \times k_g \times gain - (1 - p_s) \times k_l \times loss \quad (5)$$

However, distinct weights were used for the gain and loss components (k_g and k_l respectively), and a constant k_0 was added to capture a possible bias. The subjective probability of success (p_s) was inferred from the target size. The model assumes that participants have a representation of their precision following a Gaussian assumption, meaning that the subjective distribution of their performance could be defined by its mean (the required 1 s to reach target center) and its width (i.e., standard deviation) captured by a free parameter σ. Thus, the probability of success was the integral of this Gaussian bounded by the target window:

$$p_s = \frac{1}{\sigma\sqrt{2\pi}} \int_{1-\frac{Size}{2}}^{1+\frac{Size}{2}} e^{\frac{-(x-1)^2}{2\sigma^2}} dx \quad (6)$$

This model was inverted for participants separately using the Matlab VBA toolbox (available at https://mbb-team.github.io/VBA-toolbox/), which implements Variational Bayesian analysis under the Laplace approximation[74].

Computational analysis of stimulation effect on choice. The choice model was then run separately with data from trials with and without intracranial electrical stimulation for our 15 participants. The gain (k_g) and loss (k_l) weights of the expected utility were free to fluctuate, while all other parameters were fixed with posterior means computed with values from the previously inverted model (from the ten participants who performed an additional behavioral task).

Interaction effects of stimulation condition and clusters/subregions on posterior parameters were tested with the same glme models as the interaction effects of stimulation and clusters/subregions on choice/confidence (Eq. 2), but instead of choice/confidence, the dependent variable was the posterior parameters k_g or k_l obtained with the two trial subsets (with versus without stimulation).

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The raw behavioral data generated in this study have been deposited in OSF.

Code availability

All Matlab code necessary to reproduce our analyses is available, without restriction at https://doi.org/10.17605/OSF.IO/3JAES.

References

1. Cecchi, R. et al. Intracerebral mechanisms explaining the impact of incidental feedback on mood state and risky choice. *eLife* **11**, e72440 (2022).
2. Vinckier, F., Rigoux, L., Oudiette, D. & Pessiglione, M. Neuro-computational account of how mood fluctuations arise and affect decision making. *Nat. Commun.* **9**, 1708 (2018).
3. De Martino, B., Fleming, S. M., Garrett, N. & Dolan, R. J. Confidence in value-based choice. *Nat. Neurosci.* **16**, 105–110 (2013).
4. Lebreton, M., Abitbol, R., Daunizeau, J. & Pessiglione, M. Automatic integration of confidence in the brain valuation signal. *Nat. Neurosci.* **18**, 1159–1167 (2015).
5. Lopez-Persem, A. et al. Four core properties of the human brain valuation system demonstrated in intracranial signals. *Nat. Neurosci.* **23**, 664–675 (2020).
6. Shapiro, A. D. & Grafton, S. T. Subjective value then confidence in human ventromedial prefrontal cortex. *PLoS ONE* **15**, e0225617 (2020).
7. Hebart, M. N., Schriever, Y., Donner, T. H. & Haynes, J.-D. The relationship between perceptual decision variables and confidence in the human brain. *Cereb. Cortex* **26**, 118–130 (2016).
8. Molenberghs, P., Trautwein, F.-M., Böckler, A., Singer, T. & Kanske, P. Neural correlates of metacognitive ability and of feeling confident: a large-scale fMRI study. *Soc. Cogn. Affect. Neurosci.* **11**, 1942–1951 (2016).
9. Pereira, M. et al. Disentangling the origins of confidence in speeded perceptual judgments through multimodal imaging. *Proc. Natl Acad. Sci. USA* **117**, 8382–8390 (2020).
10. Vaccaro, A. G. & Fleming, S. M. Thinking about thinking: a coordinate-based meta-analysis of neuroimaging studies of metacognitive judgements. *Brain Neurosci. Adv.* **2**, 239821281881059 (2018).
11. Lebreton, M. et al. Two sides of the same coin: Monetary incentives concurrently improve and bias confidence judgments. *Sci. Adv.* **4**, eaaq0668 (2018).
12. Lebreton, M., Bacily, K., Palminteri, S. & Engelmann, J. B. Contextual influence on confidence judgments in human reinforcement learning. *PLOS Comput. Biol.* **15**, e1006973 (2019).
13. Howard, J. D. et al. Targeted stimulation of human orbitofrontal networks disrupts outcome-guided behavior. *Curr. Biol.* **30**, 490–498.e4 (2020).
14. Manuel, A. L., Murray, N. W. G. & Piguet, O. Transcranial direct current stimulation (tDCS) over vmPFC modulates interactions between reward and emotion in delay discounting. *Sci. Rep.* **9**, 18735 (2019).
15. Spagnolo, P. A. et al. Lack of target engagement following low-frequency deep transcranial magnetic stimulation of the anterior Insula. *Neuromodulation Technol. Neural Interface* **22**, 877–883 (2019).
16. Bechara, A., Damasio, A. R., Damasio, H. & Anderson, S. W. Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition* **50**, 7–15 (1994).
17. Bechara, A., Damasio, H., Damasio, A. R. & Lee, G. P. Different contributions of the human amygdala and ventromedial prefrontal cortex to decision-making. *J. Neurosci.* **19**, 5473–5481 (1999).
18. Clark, L. et al. Differential effects of insular and ventromedial prefrontal cortex lesions on risky decision-making. *Brain* **131**, 1311–1322 (2008).
19. Leland, J. W. & Grafman, J. Experimental tests of the somatic marker hypothesis. *Games Econ. Behav.* **52**, 386–409 (2005).

20. Reber, J. et al. Selective impairment of goal-directed decision-making following lesions to the human ventromedial prefrontal cortex. *Brain* **140**, 1743–1756 (2017).
21. Sanfey, A. G., Hastie, R., Colvin, M. K. & Grafman, J. Phineas gauged: decision-making and the human prefrontal cortex. *Neuropsychologia* **41**, 1218–1229 (2003).
22. Shiv, B., Loewenstein, G., Bechara, A., Damasio, H. & Damasio, A. R. Investment behavior and the negative side of emotion. *Psychol. Sci.* **16**, 435–439 (2005).
23. Weller, J. A., Levin, I. P., Shiv, B. & Bechara, A. Neural correlates of adaptive decision making for risky gains and losses. *Psychol. Sci.* **18**, 958–964 (2007).
24. Weller, J. A., Levin, I. P., Shiv, B. & Bechara, A. The effects of insula damage on decision-making for risky gains and losses. *Soc. Neurosci.* **4**, 347–358 (2009).
25. Duong, A. et al. Subjective states induced by intracranial electrical stimulation matches the cytoarchitectonic organization of the human insula. *Brain Stimul.* **16**, 1653–1665 (2023).
26. Yih, J., Beam, D. E., Fox, K. C. R. & Parvizi, J. Intensity of affective experience is modulated by magnitude of intracranial electrical stimulation in human orbitofrontal, cingulate and insular cortices. *Soc. Cogn. Affect. Neurosci.* **14**, 339–351 (2019).
27. Ballesta, S., Shi, W., Conen, K. E. & Padoa-Schioppa, C. Values encoded in orbitofrontal cortex are causally related to economic choices. *Nature* **588**, 450–453 (2020).
28. Ballesta, S., Shi, W. & Padoa-Schioppa, C. Orbitofrontal cortex contributes to the comparison of values underlying economic choices. *Nat. Commun.* **13**, 4405 (2022).
29. Caruana, F., Jezzini, A., Sbriscia-Fioretti, B., Rizzolatti, G. & Gallese, V. Emotional and social behaviors elicited by electrical stimulation of the insula in the macaque monkey. *Curr. Biol.* **21**, 195–199 (2011).
30. Jezzini, A., Caruana, F., Stoianov, I., Gallese, V. & Rizzolatti, G. Functional organization of the insula and inner perisylvian regions. *PNAS* **109**, 10077–10082 (2012).
31. Krolak-Salmon, P. et al. An attention modulated response to disgust in human ventral anterior insula. *Ann. Neurol.* **53**, 446–453 (2003).
32. Singh, T. D. et al. Crying with depressed affect induced by electrical stimulation of the anterior insula: a stereo EEG case study. *Epilepsy Behav. Rep.* **15**, 100421 (2021).
33. Bartolomei, F. et al. The role of the dorsal anterior insula in ecstatic sensation revealed by direct electrical brain stimulation. *Brain Stimul.* **12**, 1121–1126 (2019).
34. Nencha, U., Spinelli, L., Vulliemoz, S., Seeck, M. & Picard, F. Insular stimulation produces mental clarity and bliss. *Ann. Neurol.* **91**, 289–292 (2022).
35. Picard, F., Scavarda, D. & Bartolomei, F. Induction of a sense of bliss by electrical stimulation of the anterior insula. *Cortex* **49**, 2935–2937 (2013).
36. Destrieux, C., Fischl, B., Dale, A. & Halgren, E. Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature. *NeuroImage* **53**, 1–15 (2010).
37. Amunts, K., Mohlberg, H., Bludau, S. & Zilles, K. Julich-brain: A 3D probabilistic atlas of the human brain's cytoarchitecture. *Science* **369**, 988–992 (2020).
38. Engelmann, J. B. & Tamir, D. Individual differences in risk preference predict neural responses during financial decision-making. *Brain Res.* **1290**, 28–51 (2009).
39. Gueguen, M. C. M. et al. Anatomical dissociation of intracerebral signals for reward and punishment prediction errors in humans. *Nat. Commun.* **12**, 3344 (2021).
40. Tobler, P. N., O'Doherty, J. P., Dolan, R. J. & Schultz, W. Reward value coding distinct from risk attitude-related uncertainty coding in human reward systems. *J. Neurophysiol.* **97**, 1621–1632 (2007).
41. Venkatraman, V., Payne, J. W., Bettman, J. R., Luce, M. F. & Huettel, S. A. Separate neural mechanisms underlie choices and strategic preferences in risky decision making. *Neuron* **62**, 593–602 (2009).
42. Xue, G. et al. Functional dissociations of risk and reward processing in the medial prefrontal cortex. *Cereb. Cortex* **19**, 1019–1027 (2009).
43. Kuhnen, C. M. & Knutson, B. The neural basis of financial risk taking. *Neuron* **47**, 763–770 (2005).
44. Paulus, M. P., Rogalsky, C., Simmons, A., Feinstein, J. S. & Stein, M. B. Increased activation in the right insula during risk-taking decision making is related to harm avoidance and neuroticism. *NeuroImage* **19**, 1439–1448 (2003).
45. Pessiglione, M., Seymour, B., Flandin, G., Dolan, R. J. & Frith, C. D. Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. *Nature* **442**, 1042–1045 (2006).
46. Rolls, E. T., McCabe, C. & Redoute, J. Expected value, reward outcome, and temporal difference error representations in a probabilistic decision task. *Cereb. Cortex* **18**, 652–663 (2008).
47. Rudorf, S., Preuschoff, K. & Weber, B. Neural correlates of anticipation risk reflect risk preferences. *J. Neurosci.* **32**, 16683–16692 (2012).
48. Peng, Y. et al. Sweet and bitter taste in the brain of awake behaving animals. *Nature* **527**, 512–515 (2015).
49. Wang, L. et al. The coding of valence and identity in the mammalian taste system. *Nature* **558**, 127–131 (2018).
50. Saga, Y., Ruff, C. C. & Tremblay, L. Disturbance of approach-avoidance behaviors in non-human primates by stimulation of the limbic territories of basal ganglia and anterior insula. *Eur. J. Neurosci.* **49**, 687–700 (2019).
51. Hoven, M. et al. Motivational signals disrupt metacognitive signals in the human ventromedial prefrontal cortex. *Commun. Biol.* **5**, 1–13 (2022).
52. Ting, C.-C., Palminteri, S., Engelmann, J. B. & Lebreton, M. Robust valence-induced biases on motor response and confidence in human reinforcement learning. *Cogn. Affect Behav. Neurosci.* **20**, 1184–1199 (2020).
53. Fox, K. C. R. et al. Changes in subjective experience elicited by direct stimulation of the human orbitofrontal cortex. *Neurology* **91**, e1519–e1527 (2018).
54. Fox, K. C. R. et al. Intrinsic network architecture predicts the effects elicited by intracranial electrical stimulation of the human brain. *Nat. Hum. Behav.* **4**, 1039–1052 (2020).
55. Selimbeyoglu, A. & Parvizi, J. Electrical stimulation of the human brain: perceptual and behavioral phenomena reported in the old and new literature. *Front. Hum. Neurosci.* **4**, 46 (2010).
56. Gherman, S. & Philiastides, M. G. Human VMPFC encodes early signatures of confidence in perceptual decisions. *eLife* **7**, e38293 (2018).
57. Morales, J., Lau, H. & Fleming, S. M. Domain-general and domain-specific patterns of activity supporting metacognition in human prefrontal cortex. *J. Neurosci.* **38**, 3534–3546 (2018).
58. Ting, C.-C., Salem-Garcia, N., Palminteri, S., Engelmann, J. B. & Lebreton, M. Neural and computational underpinnings of biased confidence in human reinforcement learning. *Nat. Commun.* **14**, 6896 (2023).
59. Iravani, B. et al. Intracranial recordings of the human orbitofrontal cortical activity during self-referential episodic and valenced self-judgments. *J. Neurosci.* **44**, e1634232024 (2024).
60. Yang, Y.-P., Li, X. & Stuphorn, V. Primate anterior insular cortex represents economic decision variables proposed by prospect theory. *Nat. Commun.* **13**, 717 (2022).
61. Ghaziri, J. et al. Subcortical structural connectivity of insular subregions. *Sci. Rep.* **8**, 8596 (2018).
62. Ghaziri, J. et al. The corticocortical structural connectivity of the human insula. *Cereb. Cortex* **27**, 1216–1228 (2017).

63. Wang, R. et al. Functional connectivity gradients of the insula to different cerebral systems. *Hum. Brain Mapp.* **44**, 790–800 (2023).
64. Haber, S. N. & Knutson, B. The reward circuit: linking primate anatomy and human imaging. *Neuropsychopharmacology* **35**, 4–26 (2010).
65. Lehéricy, S. et al. Diffusion tensor fiber tracking shows distinct corticostriatal circuits in humans. *Ann. Neurol.* **55**, 522–529 (2004).
66. Price, J. L. Definition of the orbital cortex in relation to specific connections with limbic and visceral structures and other cortical regions. *Ann. N. Y. Acad. Sci.* **1121**, 54–71 (2007).
67. Lebreton, M., Jorge, S., Michel, V., Thirion, B. & Pessiglione, M. An automatic valuation system in the human brain: evidence from functional neuroimaging. *Neuron* **64**, 431–439 (2009).
68. Liu, S. & Parvizi, J. Cognitive refractory state caused by spontaneous epileptic high-frequency oscillations in the human brain. *Sci. Transl. Med.* **11**, eaax7830 (2019).
69. Deman, P. et al. IntrAnat electrodes: A free database and visualization software for intracranial electroencephalographic data processed for case and group studies. *Front. Neuroinform.* **12**, 40 (2018).
70. Auzias, G., Coulon, O. & Brovelli, A. MarsAtlas: A cortical parcellation atlas for functional mapping: MarsAtlas. *Hum. Brain Mapp.* **37**, 1573–1592 (2016).
71. Lopez-Persem, A., Verhagen, L., Amiez, C., Petrides, M. & Sallet, J. The human ventromedial prefrontal cortex: sulcal morphology and its influence on functional organization. *J. Neurosci.* **39**, 3627–3639 (2019).
72. Brainard, D. H. The psychophysics toolbox. *Spat. Vis.* **10**, 433–436 (1997).
73. Singmann, H. & Kellen, D. In *New Methods in Cognitive Psychology* 1st edn, Vol. 28 (Routledge, 2019).
74. Daunizeau, J., Adam, V. & Rigoux, L. VBA: A probabilistic treatment of nonlinear models for neurobiological and behavioural data. *PLoS Comput. Biol.* **10**, e1003441 (2014).

Acknowledgements

We would like to express our gratitude to Alizée Lopez-Persem for her invaluable assistance in identifying the vmPFC sulci on the individual anatomies. J.B. is supported by Agence Nationale de la Recherche (DECID: ANR-17-CE37-0018; CausaL: ANR-18-CE28-0016; EPICOG: ANR-22-CE17-0057).

Author contributions

J.B. conceived the study. L.M. and P.K. participated in patient recruitment and inclusion. R.C., I.R. and A.C.C. participated in data acquisition. R.C. conducted the analyses. R.C., I.R., A.C.C. and J.B. wrote the manuscript. All authors (R.C., A.C.C., I.R., L.M., P.K., M.P., J.B.) critically assessed and discussed the results, and revised and approved the manuscript at all stages.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-024-51822-8.

Correspondence and requests for materials should be addressed to Romane Cecchi or Julien Bastin.

Peer review information *Nature Communications* thanks Ming Hsu, and the other, anonymous, reviewers for their contribution to the peer review of this work. A peer review file is available.