



Preserved electrophysiological markers of confidence in schizophrenia spectrum disorder

Martin Rouy, Matthieu Roger, Dorian Goueytes, Michael Pereira, Paul Roux, Nathan Faivre

► To cite this version:

Martin Rouy, Matthieu Roger, Dorian Goueytes, Michael Pereira, Paul Roux, et al.. Preserved electrophysiological markers of confidence in schizophrenia spectrum disorder. *npj Schizophrenia*, 2023, 9, pp.12. 10.1038/s41537-023-00333-4 . hal-03826595v2

HAL Id: hal-03826595

<https://hal.science/hal-03826595v2>

Submitted on 23 Feb 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Preserved electrophysiological markers of confidence in schizophrenia spectrum disorder

Martin Rouy¹✉, Matthieu Roger², Dorian Goueytes¹, Michael Pereira¹, Paul Roux³ and Nathan Faivre¹

A large number of behavioral studies suggest that confidence judgments are impaired in schizophrenia, motivating the search for neural correlates of an underlying metacognitive impairment. Electrophysiological studies suggest that a specific evoked response potential reflecting performance monitoring, namely the error-related negativity (ERN), is blunted in schizophrenia compared to healthy controls. However, attention has recently been drawn to a potential confound in the study of metacognition, namely that lower task-performance in schizophrenia compared to healthy controls involves a decreased index of metacognitive performance (where metacognitive performance is construed as the ability to calibrate one's confidence relative to response correctness), independently of metacognitive abilities among patients. Here, we assessed how this confound might also apply to ERN-blunting in schizophrenia. We used an adaptive staircase procedure to titrate task-performance on a motion discrimination task in which participants ($N = 14$ patients and 19 controls) had to report their confidence after each trial while we recorded high density EEG. Interestingly, not only metaperceptual abilities were preserved among patients at the behavioral level, but contrary to our hypothesis, we also found no electrophysiological evidence for altered EEG markers of performance monitoring. These results bring additional evidence suggesting an unaltered ability to monitor perceptual performance on a trial by trial basis in schizophrenia.

Schizophrenia (2023)9:12; <https://doi.org/10.1038/s41537-023-00333-4>

INTRODUCTION

Schizophrenia spectrum disorder (SSD) is a mental condition with severe consequences in terms of cognitive abilities^{1,2}, social abilities^{3,4}, and more broadly on quality of life^{5,6}. For two decades, an increasing attention has been drawn to metacognitive abilities in individuals with SSD, with numerous behavioral studies suggesting an impaired ability to calibrate confidence judgments according to performance compared to healthy controls^{7,8}, paralleled with a substantial number of electrophysiological studies showing performance monitoring impairments in this population⁹. In particular, electrophysiological studies highlighted specific evoked response potentials (ERPs) which are blunted in individuals with SSD, such as the error-related negativity (ERN), the error positivity (Pe), or the feedback-related negativity (FRN)⁹. The ERN is a response-locked ERP peaking around 100 ms on frontal midline electrodes following errors^{10,11} and mostly elicited in choice reaction time tasks (e.g. Flanker task, Simon task) where participants are pressured to respond quickly (typically under 1 s), although an ERN is also found in non speeded tasks¹². The function reflected by the ERN is not clear, whether it reflects a response conflict or an error-monitoring signaling is still debated^{13–15}. It has been shown that ERN amplitude increased with confidence that one made an error¹⁶. Boldt and Yeung¹⁷ have also demonstrated a similar gradation of ERN amplitude as a function of confidence, but their multivariate analysis indicated more robust confidence modulations of Pe amplitude - a subsequent neural marker of error awareness occurring 200–300 ms after errors are committed^{9,18}. The FRN peaks 200–300 ms after negative performance feedback. Among these ERPs, the blunted ERN is considered the most robust candidate as a trait marker predictive of symptomatology^{9,19}. Here, we were interested in the link between electrophysiological markers of

performance-monitoring - such as ERN and Pe - and explicit metacognitive judgments such as the adequation between confidence ratings and task-performance. Since ERN and Pe are related to confidence but blunted among patients with SSD, we might expect a decreased ability to form relevant confident judgments (i.e. confidence judgments accurately reflecting task-performance) among patients, hence a decreased metacognitive performance.

Considering that individuals with SSD typically underperform in cognitive tasks compared to matched controls, it is important to assess if ERN-blunting simply reflects poorer behavioral performance, or if performance monitoring mechanisms are specifically impaired in SSD. In this respect, Kirschner and Klein⁹ mentioned that among the 21 reviewed studies showing a blunted ERN in individuals with SSD, 12 studies reported comparable performance between groups, while the other 9 studies reported underperformance among patients. The meta-analysis from Martin et al.¹⁹ revealed a similar pattern of results suggesting that group performance was not predictive of ERN blunting. Yet, it has been shown in healthy participants that task difficulty decreases the amplitude of the ERN²⁰, so we reasoned that *comparable* group performance, which may include differences in performance between groups that did not reach statistical significance, might still underlie subtle discrepancies between individuals that are not captured behaviorally, but that might nevertheless contaminate measures such as ERPs. In turn, little can be said about differences in ERN amplitudes between groups that behave similarly on average, but that include individuals with varying degrees of performance. Instead, we argue that performance *matching*, which uses a procedure designed to equate performance between each participant is necessary to assess the specificity of electrophysiological markers of performance monitoring in SSD. In the present

¹Université Grenoble Alpes, Université Savoie Mont Blanc, CNRS, LPNC, 38000 Grenoble, France. ²Centre d'Economie de la Sorbonne, Paris, France. ³Centre Hospitalier de Versailles, Service Hospitalo-Universitaire de Psychiatrie d'Adultes et d'Addictologie, Le Chesnay; Université Paris-Saclay; Université de Versailles Saint-Quentin-En-Yvelines; DisAP-DevPsy-CESP, INSERM UMR1018, Villejuif, France. ✉email: martin.rouy@univ-grenoble-alpes.fr

preregistered study (https://gitlab.com/nfaivre/meta_scz_public), task performance was controlled via a staircase procedure adapting the amount of sensory evidence (motion coherence) according to individual perceptual abilities. This procedure enabled us to match performance between groups and individuals, and therefore to discuss further whether the typical blunted ERN in individuals with SSD is dependent on task-performance or not.

Here we present the results from EEG data collected in patients and matched healthy controls who performed a visual motion discrimination task. Behavioral analyses of the same cohort of participants revealed preserved metacognitive abilities among individuals with SSD²¹. Building upon previous findings showing a blunted ERN in individuals with SSD^{9,19}, we conducted EEG analyses: 1) to investigate the occurrence of a blunted ERN in individuals with SSD with matched task-performance, and 2) to explore whether compensatory neural activity related to confidence (ERN-like or Pe) are found among individual with SSD, which could possibly explain why their metacognitive abilities are preserved²¹. Besides matching for performance, we also employed a paradigm which did not enforce speeded responses. By giving participants sufficient time to provide a response, we sought to quantify the electrophysiological correlates of performance monitoring without confounding our results with the detection of “slips” - a category of errors corresponding to incorrect executions of appropriate motor programs^{22,23}. Slips typically occur when participants provide a speeded response and immediately realize they pressed the wrong button, a process which differs from evaluating the probability that a decision about noisy sensory information is correct^{24,25}.

METHODS

Methods and analyses were pre-registered (NCT03140475; https://gitlab.com/nfaivre/meta_scz_public). The study was approved by the ethical committee Sud Méditerranée II (217 R01).

Participants

Twenty individuals with schizophrenia spectrum disorder (schizophrenia or schizoaffective disorder, 16 males, 4 females) and 22 healthy participants (15 males, 7 females) from the general population took part in this study. Schizophrenia and schizoaffective disorders were diagnosed by M.R. based on the Structured Clinical Interview for assessing the DSM-5 criteria. Another licensed psychiatrist (patients' treating psychiatrist) confirmed the diagnosis for each patient according to the DSM-5 criteria. The control group was screened for current or past psychiatric illness, and individuals were excluded if they met the criteria for a severe and persistent mental disorder. Five patients were excluded for having excessive artifactual EEG activity (see below), and one for having 208/300 trials with a movement onset <100 ms. Three control participants were excluded based on the following criteria: one because of a non-converging staircase, one with an estimated IQ lower than our inclusion criterion of 70, and one because no EEG data was available. In the end, the EEG analyses presented in this article were conducted on 14 individuals with SSD and 19 healthy controls.

Neuropsychological and clinical evaluation

Both individuals with schizophrenia spectrum disorders and healthy controls were evaluated on several neuropsychological domains. In particular, we assessed perceptual reasoning with the standardized score on the matrices subtest of the Wechsler Adult Intelligence Scale 4th version (WAIS-IV²⁶), verbal reasoning with the standardized score on the vocabulary subtest of WAIS-IV, working memory with the standardized score on the letter-number sequencing subtest of WAIS-IV, depressive symptoms

with the Calgary Depression Scale (CDS²⁷), cognitive insight with the composite index on the Beck Cognitive Insight Scale (BCIS²⁸). The composite index of the BCIS reflects the cognitive insight and is calculated by subtracting the score for the self-certainty scale from that of the self-reflectiveness scale. The French National Adult Reading Test (fNART²⁹) provided an estimate of pre morbid IQ.

The intensity of schizophrenia symptoms was evaluated on patients with the Positive And Negative Syndrome Scale (Kay et al.³⁰).

Experimental design

We used a visual discrimination task. Stimuli consisted of 100 moving dots within a circle (3° radius) at the center of the screen. On each trial, participants indicated whether the motion direction of the dots was to the left or to the right by reaching and clicking on one of two choice targets (3° radius circle) at the top corners of the screen with a mouse (Fig. 1A). After 6 s without response, a buzz sound rang and a message was displayed inviting the participant to respond quicker. Motion coherence was adapted at the individual level via a 1up/2down staircase procedure in order to match task-performance between groups. Following each perceptual decision, participants were asked to report their confidence about their response using a vertical visual analog scale from 0% (Sure incorrect) to 100% (Sure correct), with 50% confidence meaning “Not sure at all”. (For more details, see ref. ²¹).

In the original study²¹, we used mouse trajectories instead of button presses to investigate how the kinematics of mouse movements (velocity and acceleration) related to confidence. Here, we focused on movement initiation rather than response click as a proxy for decision time to avoid the temporal jitter due to kinematic noise in mouse trajectories (i.e. overshoots and undershoots plus small adjustments to reach the response box). We reasoned that time-locking on the initiation of the movement rather than on the response click was of particular relevance when dealing with patients with SSD, who are prone to various motor impairments, either due to medication³¹ or illness³². Movement onset was defined as the time needed from stimulus onset to reach 20% of maximum mouse velocity on each trial, from which we subtracted an arbitrary offset of five frames (~83 ms) to better capture the moment of movement initiation. Mouse movements with velocity peaks lower than 20% of the maximum velocity were discarded as non-decisional, noisy mouse movements. Visual inspection of the corresponding mouse velocity profiles showed that this procedure was effective in finding the movement onset (see Supplementary Fig. S1).

Trials with early (<100 ms) or late (>6 s) mouse movements were excluded ($6.7 \pm 7.1\%$ and $15.3 \pm 9.8\%$ of trials in controls and patients, $t = -3.06$, $p < 0.01$, $BF = 13.2$). Changes of mind occurring before the response (i.e., indicated by a change in mouse trajectory, $7.1 \pm 5.4\%$ and $8.3 \pm 8.1\%$ of trials in controls and patients, resp., $t = 0.52$, $p = 0.61$, $BF = 0.36$) or after the response (i.e., indicated by a confidence lower than 50%, $10.0 \pm 12.0\%$ and $10.7 \pm 11\%$ of trials in controls and patients, resp. $t = -0.17$, $p = 0.87$, $BF = 0.33$) were excluded, avoiding contamination from additional noise and cognitive processes. We counted as changes of mind trials in which the mouse trajectory - after having reached 20% of the total distance between the initial position of the mouse and the y-projection of the target position - crossed the midline between the two response targets.

Data analysis

Behavioral analysis. Behavioral analyses were performed using R (2020), to ensure that our behavioral conclusions in the original study²¹ were still valid on this subset of participants. In particular, we assessed whether our groups of participants were comparable both in terms of demographic and neuropsychological

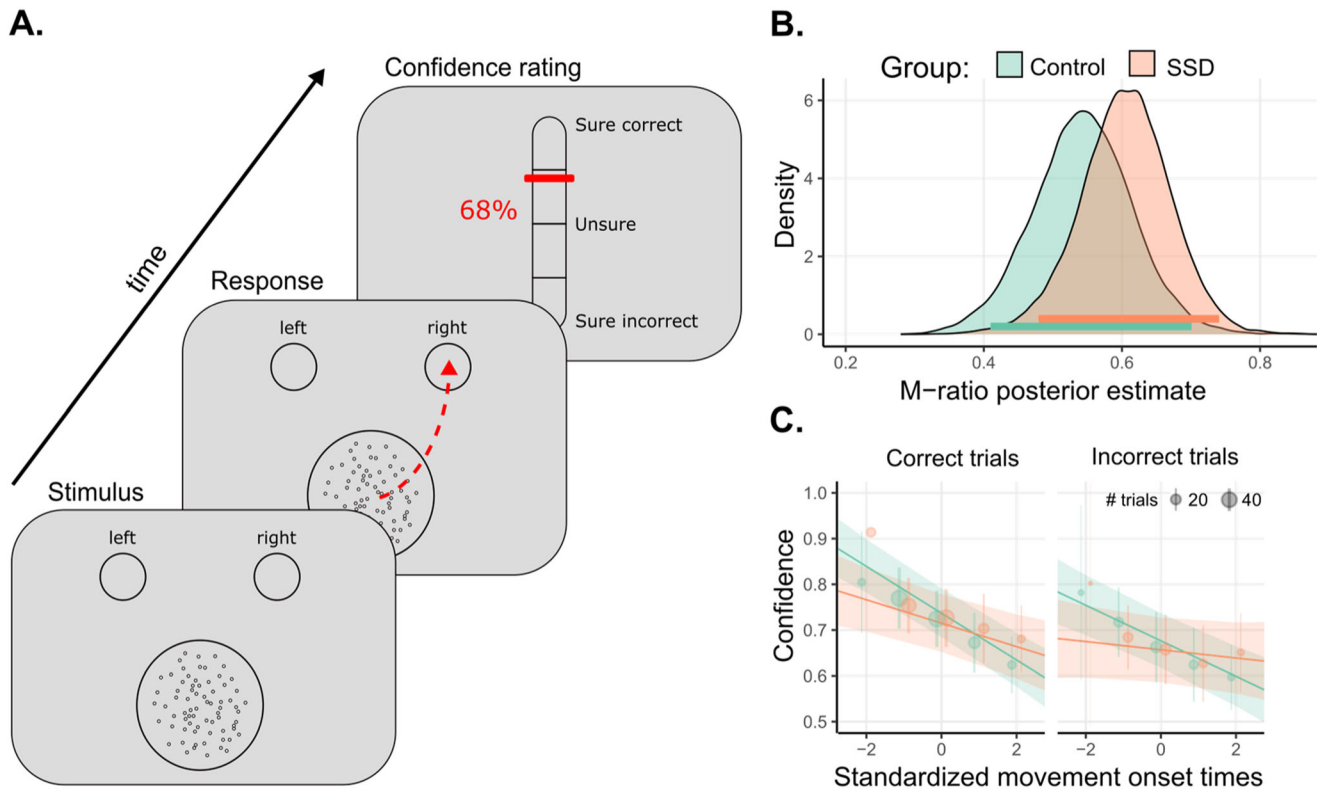


Fig. 1 Experimental task and behavioral results. **A** Experimental task: participants indicated the direction of the random-dot kinematogram by clicking in the corresponding response circle, and then estimated their confidence in their response. **B** Distribution of posterior estimates of metacognitive efficiency (M-Ratio), horizontal lines depict 95% Highest density intervals. **C** Confidence ratings regressed on standardized movement onsets, as a function of groups. Error bars indicate 95% confidence intervals. Orange: Participants with SSD; Green: Control participants. Adapted from ref. ²¹.

characteristics, and metacognitive performance (i.e. how well participants were able to calibrate their confidence judgments on performance, by computing an index of metacognitive efficiency or M-ratio³³, in a Bayesian framework³⁴ (See pre-registered plan for more details).

EEG recording and preprocessing. During the task, EEG activity was continuously recorded using a 64-channels Hlamp system (g.tec, Schiedlberg, Austria), sampled at 1200 Hz. Electrodes were positioned according to the international 10-10 system with AFz as the reference site. Impedance of electrodes was kept below 5kΩ.

Pre-processing was conducted with Matlab³⁵ scripts including functions from the EEGLAB toolbox (v2021.0, EEGLAB³⁶). EEG signal was downsampled at 128 Hz, then high-pass filtered at 0.5 Hz and low-pass filtered at 45 Hz. Bad channels were removed manually through visual inspection. Continuous EEG data were locked on mouse movement onset and epoched between -1s pre-movement to 2s post-movement. All channels were re-referenced to the common average, i.e. average over all scalp channels. Horizontal and vertical electro-oculograms were estimated by subtracting AF7 from AF8, and AFz from FPz for later identification and correction of eye-induced artifacts. At this point, noisy epochs with non stereotypical patterns of activity (as opposed to identified non-neural artifactual activity) were excluded through visual inspection. Within each individual data set, the number of components to keep for subsequent independent component analysis (ICA³⁷) was obtained through a principal component analysis, keeping only the first components contributing up to 99% of signal variance. ICA was conducted to identify artifactual components before automatic rejection using the EEG artifact detector ADJUST³⁸. Remaining noisy

epochs were excluded by visual inspection. Previously excluded channels were re-interpolated from surrounding channels using spherical splines³⁹.

To avoid spurious effects of baseline-correction, we only subtracted the average signal per subject and channel over the window from 700 to 200 ms before the movement onset. Finally, to get rid of the remaining noisy trials, we excluded the first and last percentiles of trials by individual, in terms of maximum amplitude.

Among the aforementioned exclusion of five patients for excessive artifactual EEG activity, one patient was excluded because automatic ICA rejection failed to get rid of artifactual components (170/277 trials with regular bursts of voltage amplitude remaining after ICA-based rejection). One patient was excluded because 180 trials and 9 electrodes were rejected after visual inspection. Three patients were excluded for having more than 10 channels with a variance exceeding what was found in the pool of participants by two standard deviations.

EEG data analysis. Voltage amplitude was analyzed with linear mixed-effects regressions using R⁴⁰ together with the 'lme4'⁴¹ and 'lmerTest' packages⁴². This method allows analyzing single trial data, with no averaging across conditions or participants, and no discretization of confidence ratings⁴³. Models were applied to each time sample and electrode for individual trials, to explain broadband EEG amplitude with group and correctness (resp. confidence) as fixed effects, and a random intercept per participant. False discovery rate (FDR) -correction⁴⁴ for multiple testing was applied to adjust *p*-values using the built-in R package 'stats'. Bayesian mixed-effects regressions with full random-effects structures were fitted using 'brms' R package⁴⁵.

Analyses were conducted in three steps: 1) we searched for responsive electrodes, 2) we determined the duration and amplitude of the effect at the level of the cluster of responsive electrodes, and 3) we characterized the robustness of these effects by computing evidence ratios at the cluster level (i.e. the ratio of the evidence supporting the hypothesized direction of the effect, over evidence supporting the non-hypothesized direction of the effect).

1. Search for responsive electrodes: For each time sample, electrode and independent variable of interest (i.e., correctness and confidence), we identified significant effect on voltage amplitude (FDR-corrected) within a time window from 0 to 500 ms post-movement onset with the following mixed-effects linear regressions:

$$\text{amplitude} \sim \text{correctness} * \text{group} + (1|\text{participant})$$

$$\text{amplitude} \sim \text{confidence} * \text{group} + (1|\text{participant})$$

Of note, random slopes were not added at this step as they resulted in convergence failures.

The behavioral result of a link between movement onset times and confidence ratings invited us to explore how the ERN-like ERP was modulated by movement onsets with the additional model:

$$\text{amplitude} \sim \text{movementRT} * \text{group} + (1|\text{participant})$$

where movementRT refers to movement onset times.

2. Cluster analyses: Building on the literature on ERN, we focused only on central and fronto-central electrodes as regions of interest. To avoid redundant analyses performed on each electrode separately (which are spatially close to each other), and to gain statistical power, we conducted mixed-effects linear regression restrained to the electrodes selected at step 1 which fell within our scalp regions of interest. Regressions were performed at each time sample, taking participants as random intercepts, with electrode nested within participants:

$$\text{amplitude} \sim \text{correctness} * \text{group} + (1|\text{participant}/\text{electrode}) \quad (1)$$

$$\text{amplitude} \sim \text{z_confidence} * \text{group} + \text{z_movementRT} + (1|\text{participant}/\text{electrode}) \quad (2)$$

$$\text{amplitude} \sim \text{z_movementRT} * \text{group} + \text{z_confidence} + (1|\text{participant}/\text{electrode}) \quad (3)$$

where z_confidence is standardized confidence ratings, and z_movementRT is standardized movement onset times for each participant. Since response times are known to correlate with confidence, z_movementRT was added in model (2), and z_confidence is added in model (3) as covariables of non-interest.

Of note, random slopes were not added at this step as they resulted in convergence failures. FDR-correction was applied on the resulting p-values. Only periods with significant adjacent samples extending over more than 50 ms were considered genuine effects. The voltage amplitude of the effect of correctness was computed as the average difference between correct and incorrect trials over a 50 ms window centered on the peak of the main effect of correctness. The voltage amplitude of the effect of confidence was computed as the average difference between 'Very sure' and 'Unsure' tertiles, over a 50 ms window centered on the peak of the main effect of confidence.

3. Bayesian analyses: An evidence ratio was computed on the averaged voltage amplitude over each significant spatio-

temporal cluster found in step 2:

$$\text{amplitude} \sim \text{correctness} * \text{group} + (\text{correctness}|\text{participant}/\text{electrode}) \quad (4)$$

$$\text{amplitude} \sim \text{z_confidence} * \text{group} * \text{z_movementRT} + (\text{z_confidence} * \text{z_movementRT}|\text{participant}/\text{electrode}) \quad (5)$$

Bayesian models were created in Stan computational framework (<http://mc-stan.org/>) accessed with the brms package, based on four chains of 2000 iterations including 1000 warmup samples.

For model (4) we made assumptions leading to the following prior specifications: 1) we assumed that voltage amplitude would be higher for correct versus incorrect responses with a mildly informative Gaussian prior (Mean = 1, SD = 1); 2) we assumed no difference in voltage amplitude between groups with a mildly informative Gaussian prior (Mean = 0, SD = 1); 3) we assumed a blunted ERN among patients with SSD, leading to an interaction effect of group x correctness on voltage amplitude (Gaussian prior Mean = -1, SD = 1). Other priors were by default according to the brms package in R.

For model (5) we specified the following priors: 1) we assumed that voltage amplitude would be higher (resp. lower) for higher (resp. lower) confidence ratings with a mildly informative Gaussian prior (Mean = 0.5, SD = 1); 2) we assumed no difference of voltage amplitude between groups with a mildly informative prior (Mean = 0, SD = 1); 3) Because we expected a compensatory electrophysiological signal related to confidence among patients with SSD, we assumed an increased confidence ERP among patients with SSD, leading to an interaction effect of group x confidence on voltage amplitude (Gaussian prior Mean = 0.5, SD = 1); 4) We assumed a decreasing amplitude as a function of movement onset times (Mean = -1, SD = 1); and 5) we assumed an interaction effect on amplitude between movement onset times and groups, such that movement onset times from patients with SSD would correlate less with voltage amplitude compared to healthy controls (Mean = 0.5, SD = 1). Other priors were by default according to the brms package in R.

Operationalised hypotheses were as follows:

1. The presence of an ERN is indicated by a significant main effect of correctness over frontocentral sites following movement onset,
2. The presence of ERN-blunting is indicated by a correctness x group interaction within the above mentioned cluster of electrodes, characterized by a dampened difference of voltage amplitude between correct and incorrect trials among individuals with SSD compared to healthy controls,
3. The presence of confidence-related compensatory mechanisms in patients is indicated by a confidence x group interaction, characterized by an increased difference of voltage amplitude between confidence levels in patients compared to healthy controls. We did not expect any particular localization or time-window for this effect.

All analysis scripts and data (behavioral and EEG) are publicly available (https://gitlab.com/nfaivre/meta_scz_public).

RESULTS

Demographic and neuropsychological variables

Participants with SSD and healthy controls had similar age, education level and premorbid IQ (see Table 1). However,

Table 1. Clinical and neuropsychological characteristics of our sample of participants.

	Control, M $\pm$ 95% CI (N = 19)	Schizophrenia, M $\pm$ 95% CI (N = 14)	t-statistic	p-value	Bayes factor
Age, yr	43.8 $\pm$ 5.0	38.3 $\pm$ 6.5	1.33	0.196	0.68
Education Level, yr	12.4 $\pm$ 0.4	13.0 $\pm$ 1.5	-0.69	0.502	0.43
BCIS, self-reflectiveness score	8.6 $\pm$ 1.3	16.0 $\pm$ 2.4	-5.33	<0.001	3531.51
BCIS, self-certainty score	9.5 $\pm$ 1.2	10.3 $\pm$ 2.5	-0.56	0.585	0.39
BCIS, composite score	-0.9 $\pm$ 1.7	5.7 $\pm$ 4.3	-2.81	0.012	9.10
Calgary Depression Scale, score	0.5 $\pm$ 0.5	4.4 $\pm$ 2.5	-3.01	0.009	17.84
Premordid IQ	104.2 $\pm$ 4.0	101.8 $\pm$ 4.6	0.80	0.431	0.43
WAIS Matrix subtest, score	10.1 $\pm$ 1.1	8.8 $\pm$ 1.3	1.43	0.165	0.74
WAIS letter-number sequencing subtest, score	10.0 $\pm$ 1.4	10.7 $\pm$ 1.7	-0.65	0.524	0.40
WAIS vocabulary subtest, score	8.9 $\pm$ 1.3	7.5 $\pm$ 1.2	1.62	0.115	0.88
Illness duration, yr		13.7 $\pm$ 4.2			
PANSS positive, score		16.6 $\pm$ 2.2			
PANSS negative, score		20.1 $\pm$ 2.8			
PANSS total, score		76.3 $\pm$ 8.9			
Chlorpromazine equivalent, mg		419.1 $\pm$ 159.4			

Bayes Factor >3 (resp. <0.33) indicates moderate evidence for H1 (resp. for H0).

CI Confidence Interval, BCIS Beck Cognitive Insight Scale, IQ Intelligence Quotient, WAIS Wechsler Adult Intelligence Scale, PANSS Positive And Negative Syndrome Scale.

dividuals with SSD were more depressed, with higher levels of cognitive insight (composite score) than healthy control participants, explained by a higher self-reflectiveness score.

Behavioral results

Most task-related cognitive variables were comparable between groups. In particular, both groups had similar accuracy levels (SSD: Mean = 0.73, SD = 0.03; controls: Mean = 0.74, SD = 0.02; difference between groups: $t = 1.22$, $p = 0.23$, BF = 0.61), which indicated that the staircase procedure was successful in adapting perceptual difficulty (motion variance among SSD: Mean = 1.53, SD = 0.37; controls: Mean = 2.04, SD = 0.40; difference between groups: $t = 3.75$, $p < 0.01$, BF = 35.8), with very low dispersion in task performance between participants. There was no difference in average confidence between groups (SSD: Mean = 0.71, SD = 0.12; controls: Mean = 0.71, SD = 0.14; difference between groups: $t = 0.0$, $p = 1$, BF = 0.34), indicating no confidence bias, and confidence ratings' variability was comparable between groups (SSD: Mean = 0.14, SD = 0.05; controls: Mean = 0.16, SD = 0.05; difference between groups: $t = 0.72$, $p = 0.48$, BF = 0.41). Furthermore, there was no difference in movement onsets between groups, neither for correct trials (SSD: Mean = 1.23 s, SD = 0.72; controls: Mean = 1.32 s, SD = 0.66 s, difference between groups: $t = 0.38$, $p = 0.71$, BF = 0.36) nor for incorrect trials (SSD: Mean = 1.38 s, SD = 0.84; controls: Mean = 1.53 s, SD = 0.73 s, difference between groups: $t = 0.53$, $p = 0.60$, BF = 0.38). However, there was a response side bias towards the left in patients with SSD, yet with inconclusive evidence given the BF < 3 (SSD: Mean = -0.32, SD = 0.32; controls: Mean = 0.02, SD = 0.43; difference between groups: $t = 2.56$, $p < 0.05$, BF = 2.98).

Concerning metacognitive performance, the Bayesian hierarchical model provided moderate evidence for an absence of difference between the two groups in terms of M-ratio (SSD: Mean = 0.60, 95% highest posterior density interval [95% HDI] = [0.48, 0.74]; controls: Mean = 0.55, 95% HDI = [0.41, 0.70], BF = 0.20), indicating no metacognitive deficits in our sample of participants with SSD (Fig. 1B).

Next, we investigated the relationship between confidence and movement onset. We found a negative relationship between

confidence and standardized movement onset (estimate = -0.04 [-0.05 to -0.02]; evidence ratio > 4000), indicating that confidence was higher following earlier movements. This relationship was modulated by the correctness of the responses (interaction correctness $\times$ movement onset: estimate = -0.01 [-0.02 to 0.00], evidence ratio = 124) indicating steeper slopes for correct responses, and by the group (interaction group $\times$ movement onset: estimate = 0.03 [0.01 to 0.05]; evidence ratio = 67) indicating that confidence ratings were less correlated with movement onset in participants with SSD, compared to healthy controls (Fig. 1C). Yet, the relationship between confidence and movement onset did not interact with correctness $\times$ group (interaction correctness $\times$ group $\times$ movement onset: estimate = 0.00 [-0.02, 0.01], evidence ratio = 1.83). Together, these results suggest that in this subsample of patients also, movement onsets are less predictive of confidence than among healthy controls, irrespective of correctness.

EEG analysis

Effect of correctness. Electrodes Cz, C1 and C2 responded significantly to response correctness within the 0–500 ms post-movement window, and were thus selected for further analyses. At the cluster level (Cz, C1, C2), there was a main effect of response-correctness (effect of correctness = $-0.34 \pm 0.55 \mu V$, estimate = $0.05 \pm 95\%$ CI [0.02, 0.08], evidence ratio = 165) starting 10 ms and until 330 ms after movement onset (Fig. 2). The peak of the effect occurred 266 ms after movement onset. However, there was no effect of group (estimate = $0.04 \pm 95\%$ CI [-0.13, 0.21], BF₀₁ = 10.81), nor a correctness $\times$ group interaction effect (estimate = $0.01 \pm 95\%$ CI [-0.05, 0.06], BF₀₁ = 56.47), providing evidence for an absence of blunted ERN in patients. The qualitative topographical differences seen in Fig. 2 did not reach significance.

Effect of confidence, for correct trials only. To test for a specific effect of confidence irrespective of task performance, we analyzed EEG signals as a function of confidence among correct trials only (thus including $73.7\% \pm 2.25\%$ of trials in controls and $72.7\% \pm 2.49\%$ of trials among patients in the analysis, $t = -1.10$,

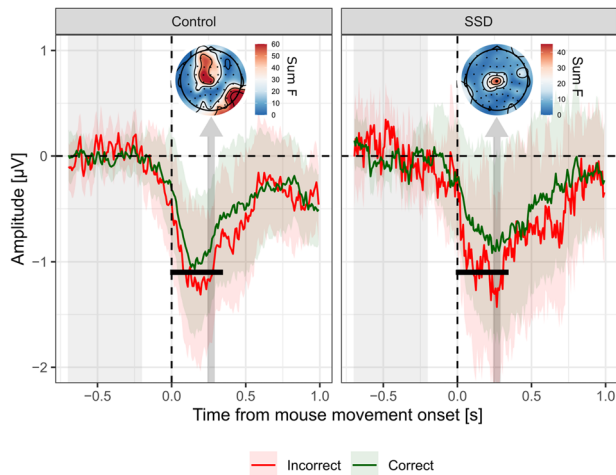


Fig. 2 ERP locked on mouse movement onset. Average signal amplitude from central electrodes (Cz, C1, C2) for incorrect responses (red), and correct responses (green). Shaded-areas represent 95% confidence intervals. Light gray shaded area on the left indicates the baseline correction window. Dark gray vertical arrows indicate the 50 ms window centered on the peak of the main effect, pointing at the corresponding topographies, scaled according to the magnitude of the main effect of correctness (summed F-values over the 50 ms window). Black horizontal lines indicate adjacent samples with a significant main effect of correctness following FDR correction.

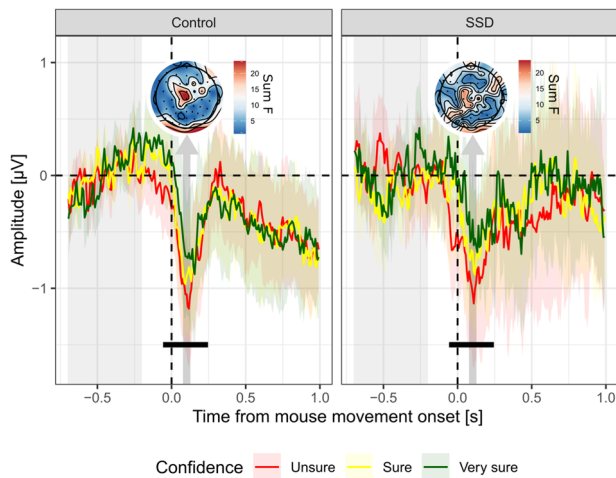


Fig. 3 ERP locked on mouse movement onset. Average signal amplitude from fronto-central electrodes (F3, F1, Fz, F2, FC3, FC1, FCz, FC2, C1, Cz, C2) on tertiles of confidence (Unsure, Sure and Very sure trials are plotted in red, yellow and green, respectively). Note that although the graphical representation is based on confidence tertiles, statistical models considered raw continuous confidence ratings. Shaded-areas represent 95% confidence intervals. Light gray shaded area on the left indicates the baseline correction window. Vertical arrows indicate the 50 ms windows centered on the peaks of the main and interaction effects (dark gray and blue, respectively), pointing at the corresponding topography, scaled according to the magnitude of the main and interaction effects of confidence (summed F-values over the 50 ms window). Black lines indicate significant adjacent samples (main and interaction effect of confidence, respectively), following FDR-correction.

$p = 0.28$, $BF_{10} = 0.54$).

Electrodes Cz, C1, C2, FCz, FC1, FC2, FC3, Fz, F1, F2, F3 responded significantly to confidence ratings within the 0–500 ms post-movement window, and were thus selected for further

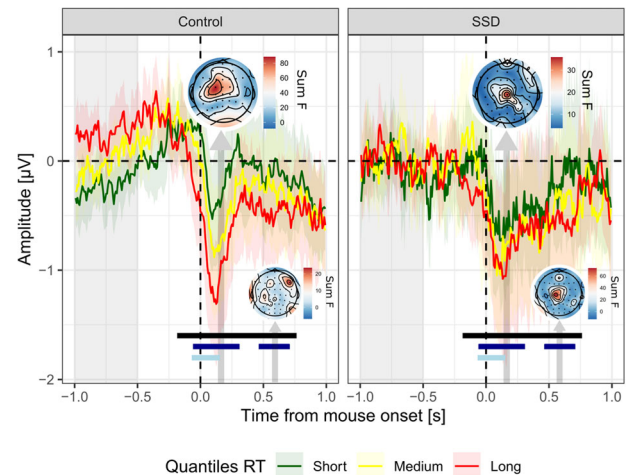


Fig. 4 ERP locked on mouse movement onset. Average signal amplitude from frontocentral electrodes (F1, Fz, FC3, FC1, FCz, FC2, C3, C1, Cz, C2) on tertiles of response times (Short, Medium and Long trials are plotted in green, yellow and red, respectively). Note that although the graphical representation is based on tertiles of movement onset times, statistical models considered continuous response times. Shaded-areas represent 95% confidence intervals. Light gray shaded area on the left indicates the baseline correction window. Vertical arrows indicate the 50 ms windows centered on the peaks of the main and interaction effects (dark gray and blue, respectively), pointing at the corresponding topography (N.B.: Occipital electrodes have been deliberately removed from the interaction topography, to focus on the activity of the ROI), scaled according to the magnitude of the main and interaction effects of response times (summed F-values over the 50 ms window). Black, dark blue, and light blue lines indicate significant adjacent samples (main and interaction effect of movement onsets, and main effect of confidence, respectively), following FDR-correction.

analyses. At the cluster level, there was a main effect of confidence peaking 102 ms after movement onset (effect of confidence = 0.41 ± 0.61 μV , estimate = $0.07 \pm 95\%$ CI $[-0.01, 0.15]$, evidence ratio = 12.5) and ranging from -40 to 227 ms after movement onset (see Fig. 3). There was no group difference in voltage amplitude (estimate = $0.02 \pm 95\%$ CI $[-0.56, 0.61]$, $BF_{01} = 3.68$), neither there was an interaction effect between $z_{\text{confidence}}$ and group within this cluster (estimate = $-0.02 \pm 95\%$ CI $[-0.16, 0.12]$, $BF_{01} = 14.8$). The qualitative topographical differences seen in Fig. 3 did not reach significance.

Overall, these EEG results are consistent with our previous behavioral results. The comparable ERN-like ERPs observed in both groups (Fig. 2) reflect the comparable correctness rates measured behaviorally, and the comparable confidence-related ERPs (Fig. 3) reflect the similarity in confidence mean and metacognitive efficiency between groups.

Effect of movement onset. As a final analysis step, we sought to investigate the relationship between response times and voltage amplitude, since response times were found to be less correlated with confidence in individuals with SSD at the behavioral level (Fig. 1.C). Electrodes Cz, C1, C2, C3, FCz, FC1, FC2, FC3, Fz, F1 responded significantly to response times and were thus selected for further analyses. At the cluster-level, there was a main effect of response times peaking 133 ms after movement onset (effect of movement onset = 0.61 ± 0.95 μV , evidence ratio = 12000) and ranging from -164 to 742 ms after movement onset (Fig. 4). Interestingly, an interaction effect between movement onset times and groups was found in two time clusters (depicted with dark blue lines in Fig. 4): the first interaction cluster ranged from -39 to 289 ms after movement onset, with moderate evidence (evidence ratio = 5.3).

This interaction cluster was characterized by a lesser relationship between EEG amplitude and movement onset times among individuals with SSD (Effect of movement onset = $0.33 \pm 0.76 \mu\text{V}$, see topography in Fig. 4), compared to control participants (Effect of movement onset = $0.67 \pm 0.71 \mu\text{V}$). The second interaction cluster ranged from 484 to 688 ms after movement onset, with moderate evidence (evidence ratio in favor of the alternative hypothesis = 7.1). This second interaction cluster described the opposite pattern compared to the first one: it was characterized by a steeper increase in voltage as a function of movement onset times among individuals with SSD (Effect of movement onset = $0.38 \pm 0.64 \mu\text{V}$, see topography in Fig. 4), compared to control participants (Effect of movement onset = $0.34 \pm 0.74 \mu\text{V}$). Furthermore, and in line with the previous analysis on confidence, there was a significant effect of the confidence covariate (depicted as light blue line in Fig. 4) from -47 to 133 ms (evidence ratio = 12.5).

DISCUSSION

In the present study, we sought to investigate EEG data recorded on patients with SSD while they performed a visual discrimination task followed by a confidence rating task²¹. Building on the literature on ERN among individuals with SSD⁹ we expected the ERN-like ERP to be blunted in the group of patients, indicating a performance monitoring deficit under matched levels of task performance between individuals of each group. Then, to make sense of the preserved metacognitive abilities at the behavioral level despite an anticipated performance monitoring deficit among patients with SSD, we hypothesized a distinctive confidence-related ERP among patients, which would constitute evidence for the existence of a compensatory mechanism helping them to provide confidence ratings that are as accurate as those provided by control participants.

We found a negative ERP over frontocentral electrodes that was larger for errors in both groups. Although the peak of this effect occurred later (266 ms) than typical ERN obtained with standard response-conflict tasks, it started 10 ms after the movement onset, consistent with the literature. This difference might be attributed to the fact that we time-locked our analysis onto the initiation of the mouse movement, which might have led to a slightly larger temporal spreading of the ERP. This has the advantage of capturing the very beginning of the decisional process instead of its end-point indicated by a button press^{46,47}. However, this may be less precise as the definition of a movement onset is temporally more ambiguous than a button press.

In line with our behavioral results, but contrary to our initial hypothesis, EEG analyses revealed evidence for unaltered neural correlates of confidence among individuals with SSD, which is consistent with the absence of a confidence bias as well as comparable variability in confidence ratings we found behaviorally between the two groups. Future research efforts with more sensitive measures and bigger sample sizes are necessary to consolidate our conclusion. We argue that such research efforts should consider matching performance experimentally between individuals of each group as comparable performance between groups may in itself not be sufficient to disambiguate ERN-blunting from poorer task performance among individuals with SSD. This argument is supported by the finding from Van der Borgh et al.²⁰ showing that ERN decreases with task-difficulty among healthy participants.

Another reason why no evidence in favor of altered correlates of confidence was found in the present study might be that the ERPs we measured are sensitive to the type of task, i.e. motion discrimination versus response-conflict tasks as commonly employed in the literature¹³. Indeed, previous studies on ERN used speeded response-conflict tasks during which participants had to suppress a prepotent response. Here, we were interested in studying

errors that arise from the slow accumulation of incorrect noisy sensory information, which are not detected as errors, by giving participants enough time to respond (6 s). It might be that the ERN-blunting is specific to fast errors committed in response-conflict tasks, which are known to involve specific mechanisms both in the memory⁴⁸ and perceptual domains^{24,25}. The difference between fast versus slow errors may be considered as involving “slips” versus “mistakes”, a terminology proposed by Reason²³. A speeded context increases the proportion of so-called “slips” - a category of errors corresponding to “incorrect executions of appropriate motor programs”²² - as opposed to “mistakes”, reflecting inaccurate intentions due to erroneous knowledge. Slips are therefore obvious errors due to participants executing the wrong motor command (pressing A and soon realizing they meant to press B). However, in a non-speeded context, errors are more likely to result from “mistakes” rather than “slips”. In sum, the ERN-blunting in speeded experiments might reflect a specific impairment regarding the monitoring of fast errors or slips, whereas the absence of ERN-blunting in the present non-speeded task might be evidence for a preserved monitoring of genuine mistakes reflected by “slow errors”. Finally, the only notable behavioral difference between the two groups was that confidence was less predicted by response times among patients with SSD (see Fig. 1D). Now extending this aspect to EEG, we found that voltage amplitudes were distinctively modulated by movement onset times among patients, within two spatio-temporal clusters (Fig. 4). At first sight, it appears striking that patients with SSD have comparable average response times, confidence levels, metacognitive abilities, and yet lower correlations between confidence ratings and response times compared to healthy controls. Zheng and colleagues⁴⁹ have also observed a lower correlation between confidence and response times among patients with SSD in a metamemory task, and their results suggest that this pattern is partly explained by a stronger reliance on previous confidence ratings (called confidence history) for the estimation of confidence in the current trial compared to healthy controls. However, when conducting the same analysis on our data, we could not find a stronger correlation between confidence and confidence history among patients compared to healthy controls, indicating that the result obtained by Zheng and colleagues did not extend to our perceptual task (Supplementary Fig. S1). Although speculative at this stage, one possibility would be that patients rely more on internal evidence than on additional cues such as response times (e.g. ref. ⁵⁰) or sensorimotor cues^{51,52} to rate their confidence.

Of note, our sample of patients with SSD was more depressed than healthy control participants. Interestingly, depression is known to enhance the amplitude of the ERN (for a review, see ref. ⁵³), and one might think that it could have compensated for the ERN blunting. Since depression is a very frequent comorbidity in schizophrenia (~50%, for a review see ref. ⁵⁴), this confound is most likely present in every ERN study including individuals with SSD, even though it is usually not discussed. Thus, depression is not sufficient to explain the absence of ERN blunting in our sample. However, depression may be sufficient to explain higher levels of self-reflectiveness (cognitive insight) among patients. Indeed, the link between depression and increased insight is now well established⁵⁵ and constitutes the “insight paradox”⁵, namely that improved insight degrades patients’ quality of life. Yet, the relationship between cognitive insight and behavioral measures of metacognition is still unclear⁵⁶. Future research experiments should assess the degree of correlation between insight and behavioral metacognition.

To sum up, we propose two alternative interpretations to explain why we found no evidence for altered neural correlates of performance monitoring among individuals with SSD: 1) Such alterations are confounded with altered task-performance in patients with SSD and are not observed anymore when task-performance is experimentally matched between individuals of each group; 2) Such alterations are specific to “fast errors” committed in response-conflict tasks, which would suggest a

specific impairment to detect fast errors or suppress prepotent responses among individuals with SSD (Morey and Rouders⁵⁷ and Addington et al.⁵⁸).

CONCLUSION

In our sample of individuals with SSD showing no metacognitive deficit at the behavioral level, we found evidence for an absence of deficits in performance-monitoring at the electrophysiological level. Larger scale studies assessing distinct types of errors while finely controlling for task performance are needed to better understand performance monitoring in SSD.

DATA AVAILABILITY

All analysis scripts and data (behavioral and EEG) are publicly available (https://gitlab.com/nfaivre/meta_scz_public).

Received: 20 October 2022; Accepted: 19 January 2023;

Published online: 23 February 2023

REFERENCES

- Gopal, Y. V. & Variend, H. First-episode schizophrenia: review of cognitive deficits and cognitive remediation. *Advan. Psychiatric Treatment* **11**, 38–44 (2005).
- Schaefer, J., Giangrande, E., Weinberger, D. R. & Dickinson, D. The global cognitive impairment in schizophrenia: consistent over decades and around the world. *Schizophr. Res.* **150**, 42–50 (2013).
- Lysaker, P. H. et al. Social cognition and metacognition in schizophrenia: evidence of their independence and linkage with outcomes. *Acta Psychiatrica Scand.* **127**, 239–247 (2013).
- Lysaker, P. H. et al. Social dysfunction in psychosis is more than a matter of misperception: advances from the study of metacognition. *Front. Psychol.* **12**, 723952 (2021).
- Davis, B. J., Lysaker, P. H., Salyers, M. P. & Minor, K. S. The insight paradox in schizophrenia: a meta-analysis of the relationship between clinical insight and quality of life. *Schizophr. Res.* **223**, 9–17 (2020).
- Hasson-Ohayon, I. et al. Metacognitive and social cognition approaches to understanding the impact of schizophrenia on social quality of life. *Schizophr. Res.* **161**, 386–391 (2015).
- Hoven, M. et al. Abnormalities of confidence in psychiatry: an overview and future perspectives. *Transl. Psychiatry* **9**, 268 (2019).
- Rouy, M., et al. Systematic review and meta-analysis of metacognitive abilities in individuals with schizophrenia spectrum disorders. *Neurosci. Biobehav. Rev.* <https://doi.org/10.1016/j.neubiorev.2021.03.017> (2021).
- Kirschner, H. & Klein, T. A. Beyond a blunted ERN - biobehavioral correlates of performance monitoring in schizophrenia. *Neurosci. Biobehav. Rev.* **133**, 104504 (2022).
- Falkenstein, M. Effects of errors in choice reaction time tasks on the ERP under focussed and divided attention in Brunia C H M, Gallard A W K, Kok A. (Eds), *Psychophysiol. Brain Res.* (pp 192–195) Tilburg, The Netherlands Tilburg University Press (1990).
- Gehring, W. J., Goss, B., Coles, M. G., Meyer, D. E. & Donchin, E. A neural system for error detection and compensation. *Psychol. Sci.* **4**, 385–390 (1993).
- Rausch, M., Zehetleitner, M., Steinhauser, M. & Maier, M. E. Cognitive modelling reveals distinct electrophysiological markers of decision confidence and error monitoring. *NeuroImage* **218**, 116963 (2020).
- Gehring, W. J., Liu, Y., Orr, J. M., & Carp, J. (2012). The error-related negativity (ERN/Ne). In *The Oxford handbook of event-related potential components* (eds. Luck, S. J. & Kappenman, E. S.) 231–291 (Oxford University Press, 2012).
- Ullsperger, M., Fischer, A. G., Nigbur, R. & Endrass, T. Neural mechanisms and temporal dynamics of performance monitoring. *Trends Cognit. Sci.* **18**, 259–267 (2014).
- Vidal, F., Burle, B. & Hasbroucq, T. Errors and action monitoring: errare humanum est sed corrigere possibile. *Front. Hum. Neurosci.* **13**, 453 (2020).
- Scheffers, M. K. & Coles, M. G. H. Performance monitoring in a confusing world: error-related brain activity, judgments of response accuracy, and types of errors. *J. Exp. Psychol. Hum. Percept. Perform.* **26**, 141–151 (2000).
- Boldt, A. & Yeung, N. Shared neural markers of decision confidence and error detection. *J. Neurosci.* **35**, 3478–3484 (2015).
- Murphy, P. R., Robertson, I. H., Allen, D., Hester, R., & O'Connell, R. G. An electrophysiological signal that precisely tracks the emergence of error awareness. *Front. Hum. Neurosci.* **6**, <https://doi.org/10.3389/fnhum.2012.00065> (2012).
- Martin, E. A. et al. ERP indices of performance monitoring and feedback processing in psychosis: a meta-analysis. *Int. J. Psychophysiol.* **132**, 365–378 (2018).
- Van der Borgh, L., Houtman, F., Burle, B. & Notebaert, W. Distinguishing the influence of task difficulty on error-related ERPs using surface Laplacian transformation. *Biological Psychol.* **115**, 78–85 (2016).
- Faivre, N. et al. Confidence in visual motion discrimination is preserved in individuals with schizophrenia. *J. Psychiatry Neurosci.* **46**, E65–E73 (2021).
- Dehaene, S., Posner, M. I. & Tucker, D. M. Localization of a neural system for error detection and compensation. *Psychol. Sci.* **5**, 303–305 (1994).
- Reason, J. T. *Human error* (Cambridge University Press, New York, 1990).
- Desender, K., Ridderinkhof, K. R. & Murphy, P. R. Understanding neural signals of post-decisional performance monitoring: an integrative review. *ELife* **10**, e67556 (2021).
- Ratcliff, R. & Rouder, J. N. Modeling response times for two-choice decisions. *Psychol. Sci.* **9**, 347–356 (1998).
- Wechsler, D., Coalson, D. L., Raiford, S. E. WAIS-IV: Wechsler adult intelligence scale (Pearson San Antonio, TX, 2008).
- Addington, D., Addington, J., Maticka-Tyndale, E. & Joyce, J. Reliability and validity of a depression rating scale for schizophrenics. *Schizophr. Res.* **6**, 201–208 (1992).
- Beck, A. A new instrument for measuring insight: the Beck cognitive insight scale. *Schizophr. Res.* **68**, 319–329 (2004).
- Nelson, H. E. & O'Connell, A. Dementia: the estimation of premorbid intelligence levels using the new adult reading test. *Cortex* **14**, 234–244 (1978).
- Kay, S. R., Fiszbein, A. & Opler, L. A. The positive and negative syndrome scale (PANSS) for schizophrenia. *Schizophr. Bull.* **13**, 261–276 (1987).
- Weiden, P. J. EPS profiles: the atypical antipsychotics: are not all the same. *J. Psychiatric Pract.* **13**, 13–24 (2007).
- Osborne, K. J., Walther, S., Shankman, S. A. & Mittal, V. A. Psychomotor slowing in schizophrenia: implications for endophenotype and biomarker development. *Biomarkers Neuropsychiatry* **2**, 100016, <https://doi.org/10.1016/j.bionps.2020.100016> (2020).
- Fleming, S. M. & Lau, H. C. How to measure metacognition. *Front. Human Neurosci.* **8**, <https://doi.org/10.3389/fnhum.2014.00443> (2014).
- Fleming, S. M. HMeta-d: hierarchical bayesian estimation of metacognitive efficiency from confidence ratings. *Neurosci. Conscious.* <https://doi.org/10.1093/nc/nix007> (2017).
- MATLAB. 9.7.0.1471314 (R2019b) (The MathWorks Inc., Natick, Massachusetts, 2019).
- Delorme, A. & Makeig, S. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J. Neurosci. Methods* **134**, 9–21 (2004).
- Delorme, A., Sejnowski, T. & Makeig, S. Enhanced detection of artifacts in EEG data using higher-order statistics and independent component analysis. *NeuroImage* **34**, 1443–1449 (2007).
- Mognon, A., Jovicich, J., Bruzzone, L. & Buiatti, M. ADJUST: an automatic EEG artifact detector based on the joint use of spatial and temporal features: automatic spatio-temporal EEG artifact detection. *Psychophysiology* **48**, 229–240 (2011).
- Perrin, F., Pernier, J., Bertrand, O. & Echallier, J. F. Spherical splines for scalp potential and current density mapping. *Electroencephalogr. Clin. Neurophysiol.* **72**, 184–187 (1989).
- R Core Team. *R: a language and environment for statistical computing* (R Foundation for Statistical Computing, Vienna, Austria, 2020), <https://www.R-project.org/>.
- Bates, D., Mächler, M., Bolker, B., & Walker, S. Fitting linear mixed-effects models using lme4. <https://doi.org/10.48550/ARXIV.1406.5823> (2014).
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. lmerTest package: tests in linear mixed effects models. *J. Stat. Softw.* **82**, <https://doi.org/10.18637/jss.v082.i13> (2017).
- Bagiella, E., Sloan, R. P. & Heitjan, D. F. Mixed-effects models in psychophysiology. *Psychophysiology* **37**, 13–20 (2000).
- Benjamini, Y. & Yekutieli, D. The control of the false discovery rate in multiple testing under dependency. *Ann. Stat.* **29**, 1165–1188 (2001).
- Bürkner, P.-C. Brms: an R package for bayesian multilevel models using stan. *J. Stat. Softw.* **80**, <https://doi.org/10.18637/jss.v080.i01> (2017).
- Pereira, M., Sobolewski, A. & Millán, J. D. R. Action monitoring cortical activity coupled to submovements. *Eneuro* **4**, ENEURO.0241-17.2017 (2017).
- Tafuro, A., Vallesi, A. & Ambrosini, E. Cognitive brakes in interference resolution: a mouse-tracking and EEG co-registration study. *Cortex* **133**, 188–200 (2020).
- Ratcliff, R. A theory of memory retrieval. *Psychol. Rev.* **85**, 59 (1978).
- Zheng, Y. et al. Atypical meta-memory evaluation strategy in schizophrenia patients. *Schizophr. Res. Cognit.* **27**, 100220 (2022).
- Kiani, R., Corthell, L. & Shadlen, M. N. Choice certainty is informed by both evidence and decision time. *Neuron* **84**, 1329–1342 (2014).
- Faivre, N. et al. Sensorimotor conflicts alter metacognitive and action monitoring. *Cortex* **124**, 224–234 (2020).

52. Pereira, M. et al. Disentangling the origins of confidence in speeded perceptual judgments through multimodal imaging. *Proc. Natl Acad. Sci.* **117**, 8382–8390 (2020).
53. Bruder, G. E., Kayser, J., & Tenke, C. E. Event-related brain potentials in depression: clinical, cognitive, and neurophysiological implications. In *The Oxford handbook of event-related potential components* (eds Luck, S. J. & Kappenman, E. S.) pp. 563–592 (Oxford University Press, 2012).
54. Buckley, P. F., Miller, B. J., Lehrer, D. S. & Castle, D. J. Psychiatric comorbidities and schizophrenia. *Schizophr. Bull.* **35**, 383–402 (2009).
55. Murri, M. B. et al. Is good insight associated with depression among patients with schizophrenia? Systematic review and meta-analysis. *Schizophr. Res.* **162**, 234–247 (2015).
56. David, A. S. Insight and psychosis: the next 30 years. *Brit. J. Psychiatry* **217**, 521–523 (2020).
57. Morey, R. D. & Rouder, J. N. BayesFactor: computation of bayes factors for common designs. R package version 0.9.12-4.2. <https://CRAN.R-project.org/package=BayesFactor> (2018).
58. Addington, D., Addington, J. & Maticka-Tyndale, E. Assessing depression in schizophrenia: the Calgary depression scale. *Brit. J. Psychiatry* **163**, 39–44 (1993).

ACKNOWLEDGEMENTS

M.P. is supported by a Postdoc.Mobility fellowship from the Swiss National Science Foundation (P400PM_199251). N.F. has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 803122). We thank Vincent de Gardelle and Jean-Christophe Vergnaud for their support during data acquisition.

AUTHOR CONTRIBUTIONS

M.Roger, P.R., and N.F. designed the study and acquired the data. M.Rouy, M.P., D.G., and N.F. analyzed the data. M.Rouy and N.F. wrote the article, which all authors reviewed. All authors approved the final version to be published and can certify that no other individuals not listed as authors have made substantial contributions to the

paper. This work has been presented at the Association for the Scientific Study of Consciousness in Amsterdam.

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/s41537-023-00333-4>.

Correspondence and requests for materials should be addressed to Martin Rouy.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023