Many Pipelines, One Dataset: Analytic Flexibility and Its Impact in EEG Research

Authors

The list of authors and their affiliations appear at the bottom of the manuscript (page 30).

Please cite the pre-print as indicated below:

Cesnaite, E., Algermissen, J., Vinding, M. C., Vitale, A., Pascarella, A., Fischer, N. L., Yang, Y.-F., Trübutschek, D., Gianelli, C., Marshall, C. R., Abbasi, O., Adamovich, T., Adel, L., Agarwal, N., Aguado-López, B., Ajmeria, U., Al, E., Aldeen, H., Alejandro, R. J., ...,Nilsonne, G., Busch, N. A. (2026). Many pipelines, one dataset: Analytic flexibility and its impact in EEG research. MetaArXiv. https://doi.org/10.31222/osf.io/wvx48_v1

Abstract

The EEGManyPipelines initiative included 396 scientists (168 teams) who independently analyzed the same dataset to test the same hypotheses, mapping the full landscape of analytical decisions. This captured an underexamined source of variability in research results: data processing choices needed to arrive at an analyzable dataset. Considering full analysis pipelines, ERP variability across teams was equivalent to variability across participants, suggesting that who analyses the data is as important as who provides it. In contrast to prior many-analyst projects, typically explaining less than 10% of variance in estimates, our granular codings of researcher choices explained 53%, 75%, and 51% of the variance in difference waves across three hypotheses, with unusual results partly attributable to deviations from the prototypical pipeline. Despite the heterogeneity in approaches and estimates, teams reached a substantial consensus on effect presence. By being transparent across all analysis stages, researchers can identify uncertainty and still draw meaningful inferences.

Scientific inquiry involves many steps between ideation and realization in the form of empirical results. Each of these steps requires researchers to make choices—often with a diversity of justifiable approaches—which can shape the ultimate outcome. Variability in the results of studies attempting to address a given research question may arise at many levels, from (a) participant characteristics and population differences, (b) features of the study design and method, (c) data preprocessing and preparation steps, and (d) the statistical inference and interpretation of results[1]. In the normal course of science, carried out by small teams of investigators, the roles of these factors are difficult to disentangle from one another. Traditional research approaches, such as meta-analysis, have limited ability to identify such sources of heterogeneity. Individual investigations differ across all of these dimensions simultaneously, data are often not available for re-analysis[2–4], and the mapping of analytic choices in published work is limited[5].

Crowd science initiatives, organized by numerous partner laboratories around the world, can start to isolate these influences by holding nearly all of them constant while allowing one to vary naturally[6–12]. Many Labs consortia repeat the same study methods across research sites, revealing surprisingly limited population heterogeneity in replication results[7,13,14]. In contrast, in many-analyst initiatives, independent researchers analyze the same dataset to answer the same set of hypotheses, rendering transparent differences in statistical approaches and empirical estimates[1,12,15–17]. With the sole exception of Botvinik-Nezer et al. (2020) in the fMRI domain, such projects have presented the crowd of participating scientists not with the raw data, but with processed and prepared data ready to analyze, potentially imposing artificial homogeneity and preventing the full diversity of research pipelines from emerging naturally. For example, Aczel et al.[18] provided analysts with datasets from the replication packages of published papers, which already had embedded within them numerous decisions by the original team. In addition, previous many-analyst projects have typically explained less than 10% of the variability in empirical estimates (e.g.,[8,12,16,19]). This limited explanatory power may partly reflect their reliance on relatively coarse codings of analysis characteristics, rather than detailed mappings of the full analytic decision process. Finally, although they recruited far more perspectives on the data than a standard small science investigation, the number of independent analyses was often small in absolute terms (e.g., 29 teams in [12], and 23 individual analysts in [19]). This calls for even larger-scale projects recruiting more analysts as well as new quantitative approaches featuring granular codings of the full stream of researcher decisions, including the precise ordering of pipeline steps.

The EEGManyPipelines project was designed to investigate preprocessing and analysis pipelines "in the wild"—that is, how researchers process and interpret the same raw data using their usual workflows[20]. Electroencephalography (EEG) is a core method in cognitive neuroscience valued for its fine-grained temporal resolution and broad accessibility. EEG analysis involves a cascade of interdependent decisions[21–23], which vary across researchers and laboratories and can affect results[24–27]. Before comparing brain responses between two conditions, researchers must decide how and in what order to filter the data, detect and remove artifacts, reference the signal, segment it into time windows, and select electrodes and time windows in which they want to test for the presence of an effect. Even preprocessing steps thus span a broad parameter space and can include epistemic as well as procedural considerations. At the same time, statistical modeling likewise differs widely across researchers[25,26].

The goals of this six-year-long initiative, involving 431 participating researchers (396 of whom analysed the initial dataset) at 358 institutions in 40 countries, were threefold. First, we sought to capture variability in the full set of researcher decisions from raw EEG data to statistical inference, as well as variability in the eventual results. As a benchmark, we quantitatively compared variability in research results attributable to analysis pipelines relative to a better known source of variability in results: the participants themselves. If cross-analyst variability matches cross-participant variability in magnitude, this confirms that researcher decisions about what to do with raw data profoundly shape research results – and thus warrant much greater consideration, transparency, and scrutiny than has been forthcoming to date. Second, we aimed to systematically map and link variability in researcher choices across their workflow to variability in their results – and by doing so address the "parsing problem"[28] that has frustrated past many-analyst projects in which most of the heterogeneity in estimates remained unaccounted for. As a

part of this effort, we quantified the extent to which specifications departed from the norm, both in the choices made and the precise ordering of those choices, and tested whether uncommon approaches account for uncommon results[29,30]. Finally, we explored whether quantitative dispersion of results can co-occur with qualitative convergence in scientific conclusions. Even if EEG research teams differ greatly in the magnitude of their estimates from the time-series data, they could ultimately agree on the presence or absence of a phenomenon. If so, scientific transparency and the increased uncertainty that can go along with it can coexist with meaningful collective conclusions.

Results

The Steering Committee of the EEGManyPipelines project selected a dataset representative of a typical EEG experiment, comprising EEG data from 33 human participants who performed a long-term memory task. We originally planned to test eight hypotheses, but narrowed our focus to the three primary ones due to time and resource constraints. We selected the three event-related potential (ERP) hypotheses (see Box 1), related to scene category (hypothesis 1), novelty (hypothesis 2), and recognition memory (hypothesis 3), including early and late ERP components, to assess how generalizable the results are across established hypotheses. These include two hypotheses with specific formulations (hypotheses 1 and 2), and one with a more open-ended formulation (hypothesis 3), which did not restrict the time window or electrode location, allowing us to explore whether hypothesis specificity influences analytic variability. Below, we report the results from 168 teams who completed the data analysis for the three hypotheses. We examine variability in analysis pipelines, variability in reported results, and the link between the two.

Hypotheses analyzed by analyst teams

1. There is an effect of scene category (i.e., a difference between images showing man-made vs. natural environments) on the amplitude of the N1 component, i.e., the first major negative EEG voltage deflection.

2. There are effects of image novelty (i.e., between images shown for the first time/new vs. repeated/old images) within the time-range 300-500 ms on EEG voltage at fronto-central channels.

3. There are effects of successful recognition of old images (i.e., a difference between old images correctly recognized as old [hits] vs. old images incorrectly judged as new [misses]) on EEG voltage at any channels, at any time.

Box 1. Hypotheses analyzed by analyst teams whose results are presented in this paper.

2.1. Quantifying divergence in approaches across teams

Several recent studies attempted to map the variability of EEG analysis practices based on published reports[26,31]. Yet there is currently no empirical reference framework for evaluating the observed diversity of real-life analysis pipelines applied to the same dataset. The EEGManyPipelines project addresses this gap by providing the first data-driven characterization of how EEG data are analyzed in practice, based on complete analysis pipelines implemented by independent research teams analyzing the same dataset. At the same time, categorizing analysis choices into discrete pipeline components necessarily involved some loss of granularity, as the richness and nuance of real-world processing decisions cannot be fully captured by predefined categories.

No two submitted pipelines were identical in their preprocessing and analysis choices. Teams used a range of common EEG analysis packages (the most common were EEGLAB, MNE-Python, and FieldTrip), and different reference schemes and high-pass filter cutoff values (Figure 1A). Pipelines included manual and automated procedures. Forty-eight teams excluded varying numbers of participants from their analysis (Supplementary Figure S1). Statistical analysis choices also varied considerably across teams (Figure 1B).

We quantified the overall agreement across teams using a normalized entropy index (NE) for each preprocessing step (constant across hypotheses) and for each analysis choice (often different across the three hypotheses). NE index values range from 0 to 1, where 0 indicates perfect agreement across teams and 1 reflects a uniform distribution of choices, corresponding to complete disagreement. For one of the 14 preprocessing steps (visual artifact-related component selection), the distribution of choices was consistent with a uniform distribution (NE=0.99, Figure 1 C). For two out of the nine analysis steps (averaging across channels, and correcting for multiple comparisons), the distribution of choices was consistent with a uniform distribution for both hypotheses 1 and 2 (both NE = 0.99). For hypothesis 2, but not others, averaging across time window bins yielded the NE of 0.99. These results suggest very high (almost maximal) variability even for these more specific hypotheses. In contrast, for hypothesis 3, no analysis step was consistent with a uniform distribution, indicating that, for this more open hypothesis, pipeline choices were actually more homogeneous. One possible explanation for this pattern is that, in the case of a more open-ended hypothesis without predefined time windows or channels, there may be greater consensus about appropriate exploratory statistical approaches, for instance, cluster statistics.

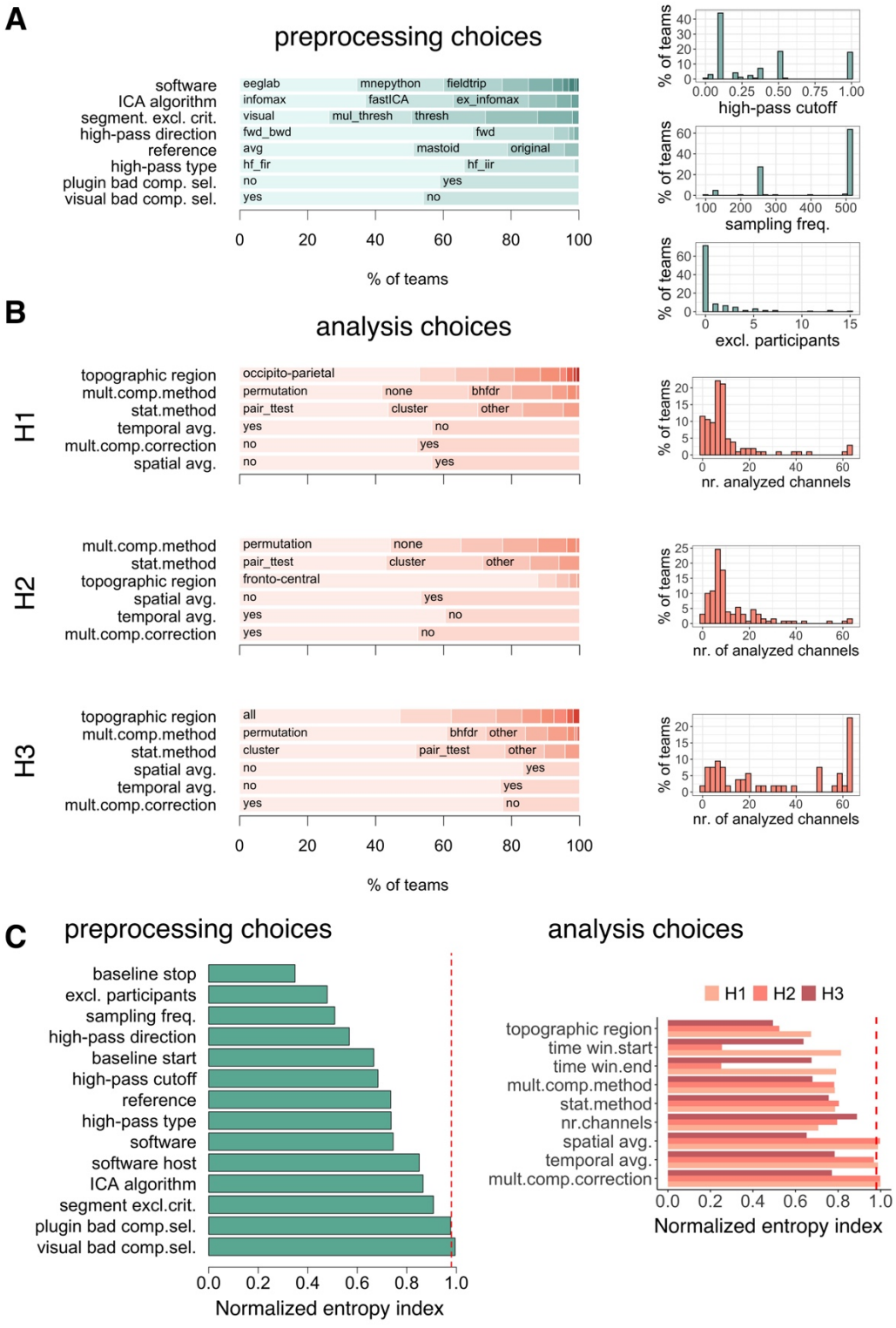

Figure 1. Variability in EEG preprocessing and analysis steps. A. *Variability in choices for preprocessing steps for categorical (left) and continuous (right) variables.* ***B.*** *Variability in choices for each analysis step, separately for the three hypotheses. Variables are ordered by the number of different choice options. The colour gradient for categorical variables reflects the frequency of each response, with lighter*

shades indicating most common choice per category. ***C.*** *Standardized entropy index for both preprocessing and analysis steps (analysis steps split across the three hypotheses, each with a different colour). The vertical red dashed line marks the normalized entropy cutoff of 0.98. Abbreviations: ICA - independent component analysis; freq. - frequency; avg - averaging, ex_infomax - extended infomax; mul_thesh - multiple threshold; fwd - forward; bwd - backward; hf - high-pass filter; fir - finite impulse response; iir - infinite impulse response; bhfdr - Benjamini-Hochberg false discovery rate; pair_ttest - paired t-test.*

To compare variability in the pipeline order, we created a transition probability matrix that contains the probability of performing step B (columns) after step A (rows, Figure 2A). We identified a team whose step sequence had the highest mean transition probability across all analyst teams and used their order of steps as the prototypical order. The prototypical order of steps was 1) high-pass filtering, 2) low-pass filtering, 3) epoching, 4) baseline correction, and 5) segment exclusion (Figure 2B). We then scored each team's preprocessing sequence based on its distance from the prototypical order of steps using the Levenshtein similarity metric[32]. The distance score reflects the minimum number of operations—insertions, deletions, or substitutions—required to transform one team's pipeline into the prototypical order: larger values indicate greater deviation from the prototype. On average, teams differed from the prototype by 4.69 operations (SD = 1.59; see Figure 2B). Given the short length of the prototypical sequence, this represents high divergence, indicating that most teams used a different ordering or diverged in the majority of preprocessing steps relative to the prototype.

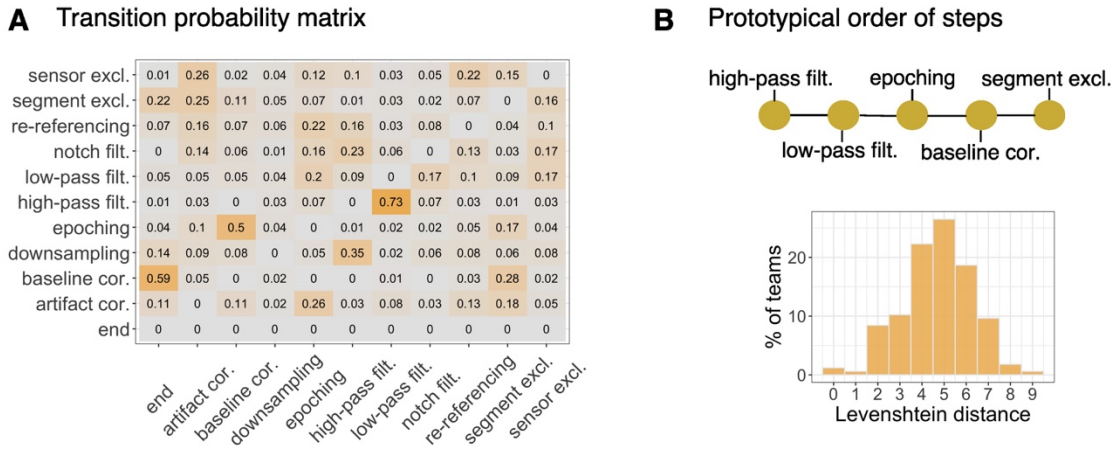

Figure 2. Variability in EEG pipeline sequences. A. *Transition probability matrix: values in each cell depict how likely it is to move from one step to another (from steps shown in rows to steps in columns). The darker colours indicate higher agreement across teams.* ***B.*** *Steps included in the prototypical pipeline sequence. The Levenshtein distance was calculated between this sequence and each team's reported pipeline to quantify deviations in processing order. Abbreviations: excl. - exclusion; filt. - filter; cor. - correction.*

2.2. Variability in hypothesis testing outcomes and effect sizes

2.2.1 Agreement on binary hypothesis outcomes masks divergence in effect sizes

Analyst teams reported binary decisions on whether each hypothesis was confirmed (Figure 3B). Confirmation rates were high for the three ERP hypotheses: on average 76.9% of teams confirmed a given hypothesis.

The binary decision whether a statistical test was significant may hide underlying heterogeneity in effect sizes. Thus, we further inspected the distribution of reported *p*-values (Figure 3C). *P*-values are hard to compare across studies due to different sample sizes. However, in our case, all analyst teams processed the same data, making the *p*-values comparable (with minor deviations in data based on which the *p*-value is based, i.e., differences in excluded epochs or participants).

The reported results and their locations in time and space varied between teams. Significant effects for hypothesis 3 were reported in a larger number of channels and broader time windows than for hypotheses 1 and 2 (Figure 3, D–E). This might be due to the open hypothesis formulation: in contrast to hypotheses 1 and 2, we did not specify the location of an effect in hypothesis 3. However, even for hypothesis 1, which addressed condition differences in the N100 ERP component, the chosen time window of interest and channel locations differed despite this specification. Hypothesis 2 predefined a time window (300–500ms) and region of interest (fronto-central electrodes). A number of teams diverged from these parameters: Nine teams (5%) examined regions beyond the fronto-central area, 38 teams (23%) did not report a specific region of interest, and 19 teams (11%) analyzed data outside the time window.

As a more direct measure of effect size, we computed ERP difference waves based on the preprocessed data submitted by each team. For each team and for each condition, we computed the mean wave across trials, subtracted the conditions from each other, and then averaged across participants. These aggregated difference waves showed substantial variability (Figure 3F).

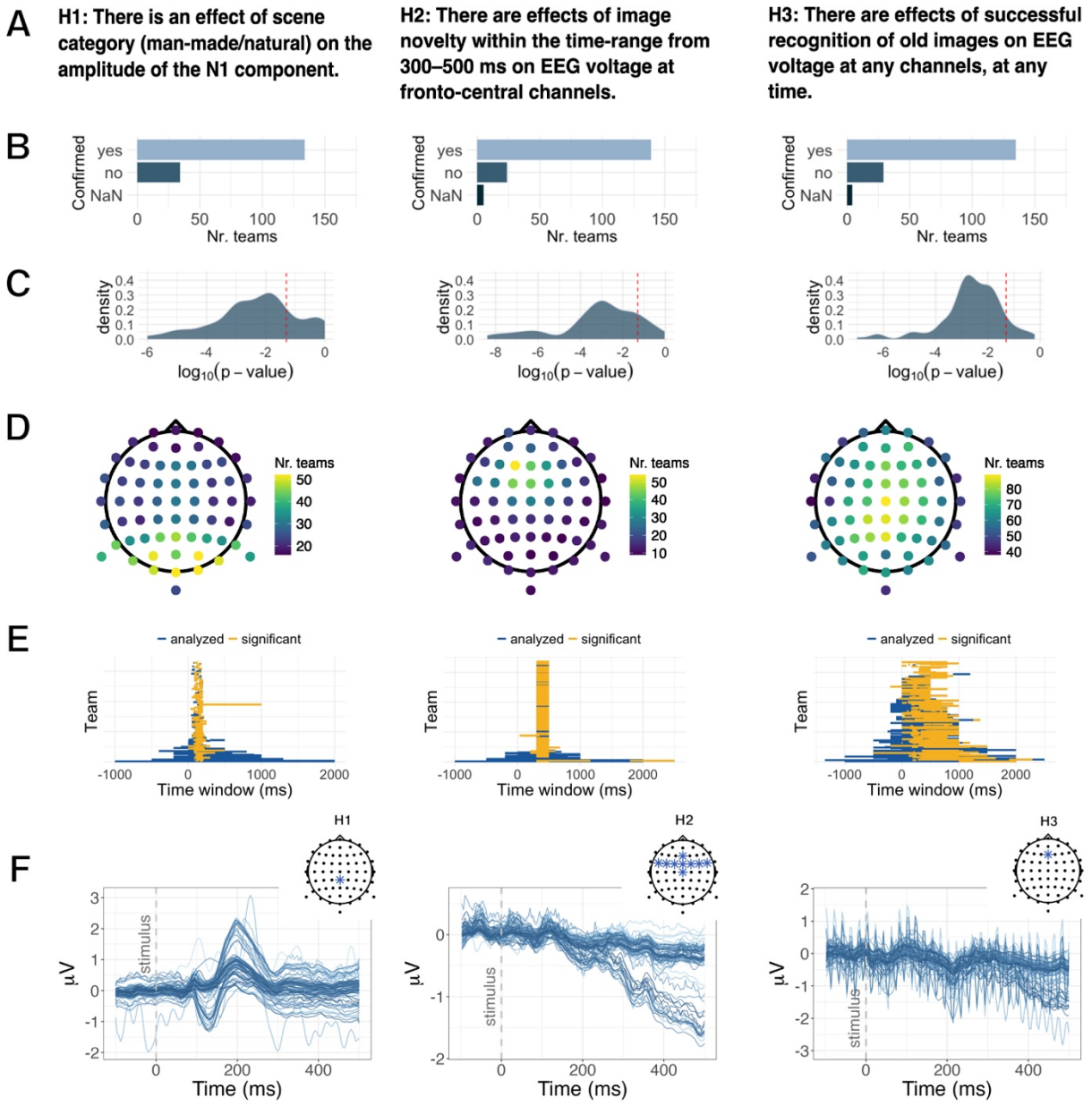

Figure 3. *Results from the testing of the three ERP-based hypotheses.* ***A****. Hypothesis formulation.* ***B****. Hypothesis confirmation rates across teams. The 'NaN' column indicates the number of teams that did not indicate whether a hypothesis was confirmed.* ***C****. Density plots of* log_{10}*transformed reported p-values. The red dashed line represents a p-value threshold of 0.05.* ***D****. Location of the reported channels that showed a significant effect.* ***E****. Analyzed time windows and time windows where a significant relationship was reported for every team.* ***F****. Difference waves and corresponding EEG channels used for ERP difference calculations. One line corresponds to one team. Note that some difference waves contain a prominent high-frequency component, which might be a result of an absence of a notch filter or a small number of epochs for one of the conditions.*

2.2.2 Variability in results across teams was comparable to variability across research participants

We compared variability in difference waves to a familiar benchmark in EEG research: variability across participants, typically accounted for in second-level analysis (denominator of statistical tests). Comparing the magnitudes of these two sources of variance is informative because EEG researchers have strong intuitions about between-participants variance, which serves as an important parameter in most statistical tests, while much less is known about the between-pipelines variance.

First, we computed the average ERPs across all trials (independent of conditions) for each participant and team. We then computed both (a) the average for each team, allowing us to quantify variability across participants (Figure 4, left column), and (b) the average for each participant, allowing us to quantify variability across teams (Figure 4, middle column). In addition, we applied the same procedure to the difference waves for hypothesis 1 (splitting trials by condition and subtracting conditions from each other).

Grand-average ERP and difference waves varied across both participants and teams (Figure 4), with variability on a comparable scale across the two levels of analysis. Especially for late-latency components of the hypothesis 1 difference wave (components above 300 ms post-stimulus), the across-teams variance was lower than the across-participants variance (mean team variance = 0.19; mean participant variance = 0.34).

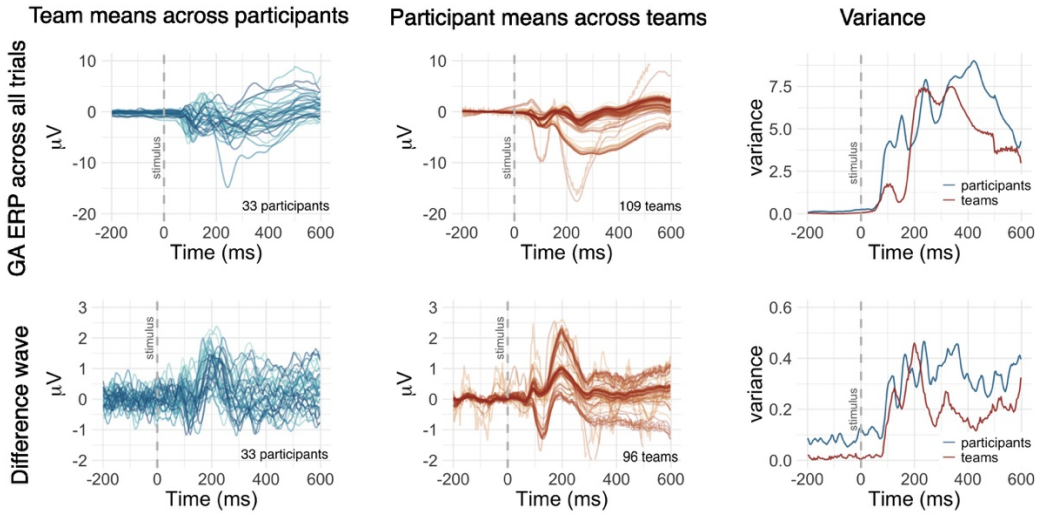

Figure 4. Grand average (GA) event-related potentials (ERP) (top) and hypothesis 1 difference wave (bottom) at the CPz channel across teams and across participants. *The left column shows the variability across participants (each line is data from one participant averaged across teams). The middle column shows the variability across teams (each line is data from one team averaged across participants). The right panel shows the variance calculated across teams and across participants. For visualisation purposes we baseline corrected the GA ERPs and difference waves prior to variance analyses.*

2.3. Association between results and analysis choices

2.3.1 Few pipeline choices predicted p-value size; most of the variance remained unaccounted for

We used regression models to test whether specific pipeline choices predicted the results. Given that linear models assume normally distributed residuals, and that the distribution of *p*-values was strongly right-skewed, we transformed *p*-values into *z*-values using the inverse normal quantile function and used this metric as a dependent variable in linear models and pipeline choices as independent variables. Due to the high number of pipeline steps, we used an automated model comparison tool, which performed stepwise forward-and-backwards variable selection based on Bayesian Information Criterion (BIC, for more details see Methods section 4.7).

Hypothesis 1. A model including the number of channels (β = -0.31 ± 0.12, CI = [-0.55, -0.07], p = 0.011), baseline time window length (β = 0.35 ± 0.14, CI = [0.07, 0.63], p = 0.015), and correction for multiple comparisons (a binary variable; β = −0.89 ± 0.26,CI = [-1.40, -0.39], p < 0.001) yielded the lowest BIC (BIC = 158.28). This model explained R^2 = 0.13 of the variance in p-values for hypothesis 1. These results suggest that a larger number of analyzed channels, a shorter baseline window length, and correction for multiple comparisons are linked to smaller p-values (Figure S2). The counterintuitive finding that correction for multiple comparisons is linked to smaller p-values might be explained by the observation that most of the teams used permutation tests as a method of multiple comparison correction for hypothesis 1 (see Figure 1). Note that these coefficient estimates and associated *p*-values are conditional on the BIC-selected model and are interpreted as descriptive indicators of relative predictor importance rather than formal significance tests.

We also tested interactions between the baseline window length and correction for multiple comparisons, and between the number of channels and correction for multiple comparisons. The interaction between the baseline window length and correction for multiple comparisons was statistically significant (β = -0.95 ± 0.35, CI = [-1.64, -0.26], *p* = 0.007, Figure S2). This means that the baseline window length was more strongly positively associated with *p*-values in teams that did *not* correct for multiple comparisons relative to those that did. It may indicate that in the absence of multiple comparisons, the larger baseline window length may inflate the likelihood of a false-negative statistical result.

Hypothesis 2. Stepwise model selection using the BIC did not identify any subset of predictors that improved fit compared to the null model. This suggests that none of the analysis steps significantly explain variation in the *p*-value size.

Hypothesis 3. Stepwise model selection using the BIC identified a single predictor: averaging across the spatial domain (a binary predictor; β = -0.54 ± 0.23, CI = [-0.98, -0.09], *p* = 0.019*;* BIC = –21.18, see Figure S3). This model explained 4% of the variance in *p*-values for hypothesis 3. This finding suggests that teams that averaged across the spatial domain tended to report lower *p*-values.

In summary, associations between analysis choices and team-reported *p*-values were inconsistent across hypotheses and were overall sparse; most preprocessing and analysis choices were not systematically or strongly related to the obtained *p*-value. This dependent variable, notably utilized as the primary outcome in some prior many analyst projects (e.g, [19]), may have contributed to the limited variance successfully explained.

2.3.2 Reference schemes, filtering, and component selection accounted for substantial variance in ERP difference waves

To test how processing pipelines affect the actual ERPs, we compared the preprocessed EEG data submitted by the analyst teams. Effects derived from these data are directly comparable across teams, as they are not contingent on team-specific statistical modelling or inference choices. We obtained these effect sizes by creating an ERP difference wave (see Methods section 4.5.2). A model comparison with BIC was used to select the optimal model predicting the difference wave values from preprocessing steps.

Hypothesis 1. We averaged the difference wave from 80–120 ms (a prototypical time window for N100 effects) at the CPz channel for each team. A model including the reference channel and the use of a bad-component selection plugin (a binary choice) as predictors yielded the lowest BIC (BIC = -305.33). The model explained 53% of the variance in the size of the difference waves. Depending on choice of the reference scheme, the sign of the N100 difference wave flipped: While it was positive in teams who used an average or a mastoid reference (n = 74), suggesting larger ERP amplitudes for man-made than natural stimuli, it was negative in teams who used a POz reference (n = 25), suggesting larger ERP amplitudes for natural than man-made stimuli categories. Difference waves were significantly smaller for teams with POz as reference compared to the teams with an average reference ($\beta = -0.45 \pm 0.05$, CI = [-0.55, -0.34], $p < 0.001$, Figure 5). Furthermore, the difference wave was slightly larger for teams with a mastoid reference than teams with an average reference ($\beta = 0.07 \pm 0.04$, CI = [-0.018, 0.16], $p = 0.118$).

In addition, teams that used a plugin-based implementation to select bad components had larger ERP difference waves for hypothesis 1 testing as compared to teams that did not ($\beta = 0.09 \pm 0.04$, CI = [0.01, 0.17], $p = 0.031$, Figure 5). Of the 70 teams that used a plugin-based implementation, 34 used ICLabel[33].

Hypothesis 2. We averaged the difference waves from 300–500 ms (as indicated in the hypothesis formulation) at a cluster of fronto-central channels. The model including reference channel and high-pass filter cutoff value choice as predictors yielded the lowest BIC (BIC = -259.43). This model explained 75% of the variance in the size of the difference waves. Teams who used a mastoid reference found a larger difference between the two conditions than teams who used an average reference ($\beta = -0.7 \pm 0.07$, CI = [-0.84, -0.56], $p < 0.001$) or a POz reference (POz compared to average reference: $\beta = -0.03 \pm 0.03$, CI = [-0.09, 0.02], $p = 0.24$, Figure 5). Furthermore, teams with a lower high-pass filter cutoff found a larger difference than teams with a higher cutoff ($\beta = 0.07 \pm 0.01$, CI = [0.03, 0.10], $p < 0.001$, Figure 5).

Hypothesis 3. We averaged the difference waves from 300–500 ms (a prototypical window of the late positive complex). Since hypothesis 3 did not specify a channel, a component, or a time

window for the analyses (and therefore is the most open of the hypotheses that were tested in this paper), we selected two locations: the Fz and Pz channels. Since the effects were comparable, we report the results obtained from the Fz channel in the text and the results for Pz in Table S7. The model containing the choice of reference channel and high-pass filter cutoff as predictors yielded the lowest BIC (BIC = -241.9). The model explained 51% of the variance in the size of the difference waves. Teams who used a mastoid reference found a larger difference wave compared to teams who used an average reference (β = -0.45 ± 0.07, CI = [-0.59, -0.32], p < 0.001) or POz reference (β = -0.10 ± 0.05, CI = [-0.20, 0.001], *p* = 0.05). Furthermore, teams with a lower high-pass filter cutoff found a larger difference compared to teams with a higher cutoff (β = 0.06 ± 0.03, CI = [0.002, 0.11], *p* = 0.043, Figure 5).

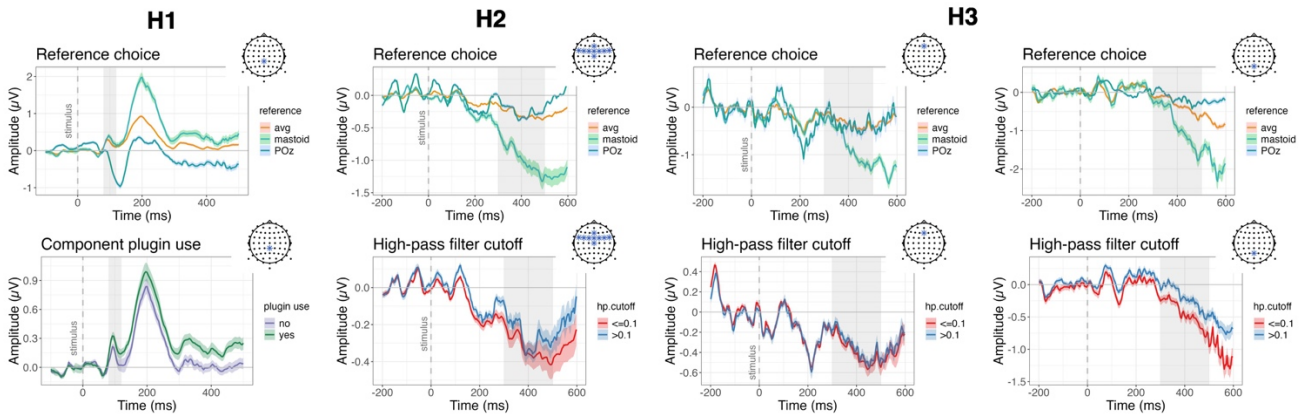

Figure 5. Median difference waves for the three hypotheses for different pipeline choices. *The difference wave is plotted for the reference choice and component plugin use for hypothesis 1 (H1). The difference waves of hypotheses 2 (H2) and 3 (H3) are plotted for the reference choice and high-pass filter cutoff values at different channels. The scalp topographies show the locations of channels from which the difference waves were extracted. The shaded grey areas represent the time window over which the difference wave was averaged. Note that the high-pass filter cutoff value is grouped into two classes for visualisation. Abbreviations: avg - average; hp. - high-pass filter.*

2.3.3 More unusual pipelines are associated with flatter difference waves for hypothesis 2

To quantify the unusualness of a team's pipeline, we calculated (a) the step-order prototypicality by using Levenshtein distance score (described in 2.1 and 4.5.1), and (b) the parameter-choice prototypicality, defined as the mean probability of the specific choices made by a team across preprocessing steps.

We then correlated the median absolute amplitude of the difference waves, reflecting the unusualness of the preprocessed EEG data (see Methods section 4.6), with the two metrics quantifying the unusualness of the corresponding analysis pipelines (step-order prototypicality and parameter-choice prototypicality indices). We observed directionally negative correlations between the difference waves and the Levenshtein distance scores regarding step-order

prototypicality (Figure 6), indicating that pipelines closer to the prototypical sequence tended to yield larger amplitudes, whereas more atypical step orders were associated with flatter difference waves. This relationship reached statistical significance only for hypothesis 2 (ρ = -0.28, p = 0.009). For hypotheses 1 and 3, the association was in the same direction but did not reach significance (ρ = -0.09, p = 0.38, and ρ = -0.10, p = 0.35, respectively). For the parameter-choice prototypicality index, we found no statistically significant relationships with the unusualness of the difference waves for any hypothesis.

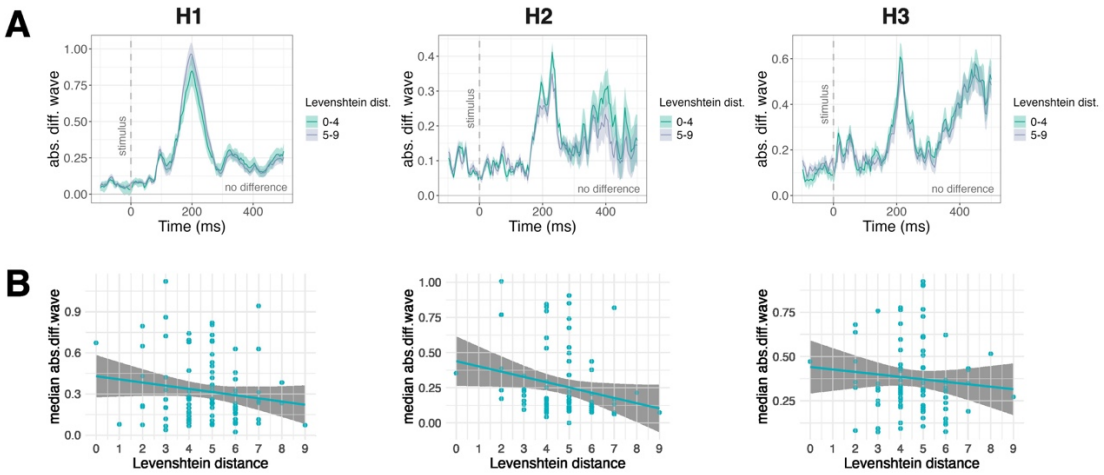

Figure 6. Associations between the unusualness of the order of steps and difference waves. A. *Absolute amplitudes of difference waves of three hypotheses, color-coded based on the team's Levenshtein distance score (denoting the distance to the prototypical order of steps, also see Figure 2). The shaded area around the lines shows the standard error. For hypothesis 1 we used channel CPz, and for hypotheses 2 and 3 we used channel Fz.* ***B.*** *Correlations between the median absolute difference wave amplitude in the time window between 100-600ms post stimulus and Levenshtein distance.*

Discussion

In the EEGManyPipelines project, 168 teams totalling 396 individual researchers analyzed the same dataset, enabling a meta-scientific investigation into how differences in analysis pipelines shape research outcomes. All teams tested the same hypotheses, thereby eliminating several major sources of variability when comparing across different research investigations. Despite this, pipelines differed in included channels, time windows, statistical models, and reported metrics. This large-scale, coordinated effort provides the first systematic quantification of differences in EEG preprocessing and analysis choices, variability in outcomes, and how these relate to one another, sampling realistic pipelines from a large group of active EEG researchers. At the same

time, it provides insights for any scientific discipline and investigation that grapples with ambiguities in complex data.

Critically, our crowd of analysts started from the raw observations, allowing us to capture not only heterogeneity in analytic specifications, but also in data preprocessing steps (see also[15]). The latter remains an almost entirely shrouded aspect of researcher decision-making, with publicly posted and shared datasets taking the form of processed data in the vast majority of cases. Examining full research pipelines, including the precise ordering of pipeline steps, is especially important in fields such as EEG research, where investigators regularly work with large amounts of nested time-series data. For this purpose, our crowd of analysts submitted preprocessed data, which allowed us to quantify variation in effect sizes attributable to the subsequently chosen data-handling approach. Our findings reveal that choices regarding preprocessing parameters, such as the reference channel, high-pass filter cutoff, and the way components associated with noise are selected, predicted difference wave magnitude (i.e., the ERP difference between two conditions). Comparing different sources of variability in ERP difference waves, cross-analyst variability matched cross-participant variability in magnitude in this data, a descriptive benchmark for variability in results attributable to researcher choices that we introduce into the literature here. It is important to note that this comparison may depend heavily on several factors, including data quality, the nature of the research question, and specific analytic contexts. However, at least in this setting, who analyses the data is just as important as who provided the data—a testament to the power of researchers' degrees of freedom to shape the outcomes of scientific investigations.

Despite variability in the difference waves, results for the three ERP hypotheses revealed a majority agreement across teams: 76.9% of teams on average agreed on the presence of a statistically significant effect. The hypotheses in question targeted well-established ERP effects related to scene category (hypothesis 1), novelty, and recognition memory (hypotheses 2 and 3)[34,35]. Notably, agreement rates were comparable across hypotheses, despite differences in how specifically the ERP effects were defined, suggesting that the level of constraint in the hypothesis did not substantially influence overall consensus. It is conceivable that agreement across teams would look different in projects addressing novel or controversial hypotheses or relying on less established experimental paradigms. Regardless, simple binary judgments of whether an ERP hypothesis was confirmed or not camouflaged divergence in the underlying data patterns. Difference waves not only varied in magnitude, but in some cases reversed in polarity. For example, for hypothesis 1, while most teams identified a positive-going ERP for pictures of man-made scenes, twenty-five teams found a directionally opposite pattern, with a larger amplitude for pictures depicting natural scenes. Thus, two papers addressing the same question and analyzing the same dataset could plausibly report ERP difference waves that lead to opposing conclusions. To facilitate collectively adjudicating such issues, researchers should share their underlying preprocessed data and code in an accessible format post-publication.

Similar to past many-analyst projects, which typically explain under 10% of the dispersion in approaches and estimates (e.g.,[16,19]), we found that variation in results based on transformed *p*-values could hardly be explained by pipeline choices. Because all teams analyzed the same raw data, the full variability in reported effects necessarily arose from differences in analysis pipelines. Rather, the low explained variance may reflect the coarse nature of the outcome measure and

the exclusion of some idiosyncratic analytic decisions, such as trial or participant rejection criteria. This underscores the importance of not only providing granular codings of researcher approaches but also precise measurement of outcomes to solving the parsing problem[28] posed by radical dispersion in estimates.

Indeed, we found that pipeline choices for the three hypotheses could be used to quantitatively explain the majority of the variability in ERP difference waves derived from analysts' submitted preprocessed data—53%, 75%, and 51% of the variance, respectively. The reference channel emerged as the pipeline choice that explained the largest proportion of variance in all difference waves[36,37]. The influence of the specific reference location was hypothesis-dependent. For hypothesis 1, teams using POz as the reference channel produced numerically larger N100 difference waves than those using mastoid or average references. However, the direction of the effect was reversed for POz-referenced teams: they observed larger ERP amplitudes for natural stimuli than for man-made stimuli, whereas teams using other reference schemes showed the opposite pattern. Given that ERP effects in experiments using visual stimuli typically have a posterior topography, and that most researchers chose to re-reference the data, keeping the original reference might be considered a questionable choice. However, note that participating EEG laboratories were asked to use their standard pipeline, that keeping the original reference resulted in a plausible difference wave, and that the choice of this particular reference scheme could have gone unnoticed in regular journal publications[26]. Still, the question of defensibility or quality of different pipeline choices remains an important one. Alternative branches of choices should be weighted by methodological quality and parsed with regard to their theoretical implications[29,30].

For hypotheses 2 and 3, the high-pass filter cutoff value was also a significant predictor of the difference wave magnitude: higher cutoff frequencies were associated with smaller difference waves, suggesting that the distinction between the two conditions became less pronounced. This observation corroborates previous findings that higher filter cutoff can compromise ERP waves[38–40]. In addition, the use of an automated plugin to detect artifact-containing components was linked to the size of the difference wave of hypothesis 1: teams that used the plugin-based implementation reported larger ERP difference waves as compared to teams that did not use the plugin. The effectiveness of independent component analysis critically depends on the quality of source separation, which can be influenced by upstream preprocessing and data-cleaning practices. However, the automated independent component classifiers, which are trained on expert-labelled component examples, may outperform visual inspection conducted by a single analyst, because they provide more standardized and experience-informed artifact identification. This conclusion would have to be further corroborated by projects using simulated data sets (with known artifact components) and by explicitly evaluating automated plugins in their ability to filter out artifactual components.

The complex, multi-purpose nature of EEG preprocessing and analysis pipelines means that many variables are needed to describe heterogeneity in these pipelines. Testing for the moderating effects of a specific analytic choice across specifications may mask the extent to which a combined set and ordering of choices are normative vs. deviates from common practice. Therefore, we defined the step-order prototypicality and parameter-choice prototypicality, and

quantified the deviation of each team's pipeline from these prototypes. Linking the step-order prototypicality metric to the difference wave, we observed that teams with more unusual pipelines (i.e., larger deviation from the prototypical order of steps) tended to observe weaker effects (i.e., smaller deviations of the difference waves from zero) for one of three hypotheses. Partly supporting some prior critiques of crowd science[29,30], use of prototypicality metrics confirms that unusual pipelines can contribute to unusual results. However, the prototypical pipeline does not necessarily reflect the best pipeline to use in a given analysis. Future research using simulated data in which the true underlying effect size is known is needed to determine whether an overall assessment of a pipeline as "typical" vs. "atypical" is a good first indicator for its validity.

It is also important to emphasize that the generalizability of the conclusions from this project could be limited to the shared particular dataset. Despite this caveat, we believe this first EEGManyPipelines initiative represents a sufficiently striking "existence proof" regarding the potential impact of researcher choices to hold some implications for routine EEG research practice. First, analytic pipelines should be conceptualized as central components of study design. Second, key preprocessing decisions, particularly reference schemes and filter settings, should be subjected to robustness checks. Third, reliance on binary statistical judgments may obscure meaningful variability in underlying waveform patterns; complementing such judgments with continuous effect measures may therefore improve transparency and interpretability.

For the broader scientific community, our findings indicate that efforts to meaningfully explain variability in results and reach collective conclusions should take into account full analysis pipelines. This necessarily includes data preprocessing steps and their order, currently left shrouded as a matter of routine research practice. Achieving this level of transparency requires the widespread sharing of raw and preprocessed data, code, and sufficiently detailed documentation—ideally through machine-readable, structured reporting tools such as ARTEM-IS[41]. To streamline this documentation process, analysis software should actively support transparent reporting by automatically generating ready-to-use, guideline-compliant method descriptions that accurately reflect the analytic procedures employed. Finally, we recommend that individual research teams routinely assess the robustness of their findings to variation in key data preprocessing and analysis decisions (e.g., reference schemes, filter cutoff values, tools for artifact correction), and transparently report how such choices influence their results.

Methods

The steering committee of the EMP project was formed online, in a bottom-up grass-roots fashion[20], in the summer of 2020. They selected a dataset to be analyzed, specified target hypotheses that could be tested given the data, and recruited teams of analysts with certain stipulations[42]. Analyst teams were asked to test hypotheses using their laboratory's typical analysis pipeline on the provided dataset. The steering committee analyzed the collective meta-scientific data to map variability in pipelines, heterogeneity in results, and possible links between analytic decisions and results. In recognition of their contributions, all participating analysts were offered co-authorship on the resulting many-analyst journal articles.

Data was submitted in different formats. The steering committee transformed it into a consistent format for easier data analysis. The steering committee also checked the submitted descriptions of teams' analysis pipelines against the submitted code, identified discrepancies, and resolved any such discrepancies (relying on the code). All processed data, scripts, and response data provided by the analysts are available at https://brainlife.io/project/6863bf5c1521e536327bfea7 and described in [43].

The study was approved by the ethics committee of the Faculty of Psychology and Sports Sciences at the University of Münster, Germany (#2021-06-NB and #2021-52-NB-FA).

4.1 Enrolment and characteristics of analyst teams

Analysts were recruited via mailing lists, social media, and word-of-mouth from April to October 2021. A total of 327 teams registered for the project before the start of the analysis phase. Teams could consist of one to three researchers, but were required to feature at least one member who had previously published an EEG study in a peer-reviewed journal to be eligible. The final EEGManyPipelines sample comprised 168 teams who submitted complete analyses and results for all eight hypotheses. These teams included 396 individual researchers (mean age 33.7 ± 7.5 years; 150 women, 223 men, 2 gender-diverse, 21 undisclosed), representing diverse nationalities, career stages, and subfields of neuroscience (including memory, attention, and related areas, see Figure 7). More details on sample characteristics and dropout rates at different project stages are provided in the accompanying data descriptor paper[43].

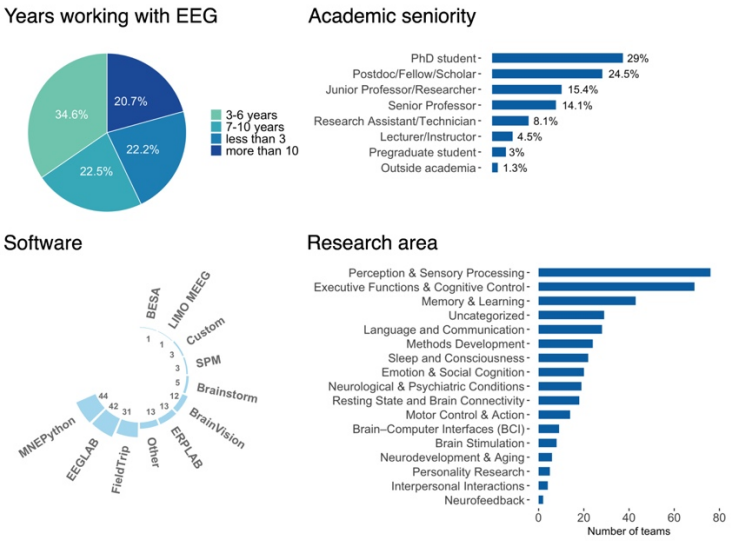

Figure 7. Demographic profile of analyst teams*. Details include years of experience with EEG, academic seniority, preferred software for data analysis, and primary research areas.*

4.2 EEG dataset

EEG data were recorded from 33 participants (22 female; 4 left-handed; mean age: 27 years) performing a continuous recognition memory task[44,45]. Participants viewed a stream of scene images (man-made: highways, buildings; natural: forests, beaches), some of which repeated, and indicated whether each image was “new” (first presentation) or “old” (repeated) by lifting a finger from one of two response keys (left/right CTRL). Response category assignments (old/new) were counterbalanced across participants. Each image was displayed for 0.5 seconds on a black background, followed by a blank screen presented until a response was generated. Feedback was given via a green fixation cross for correct responses and a red cross for incorrect responses, each displayed for 0.2 seconds. This dataset was selected because it comprises multiple experimental conditions, making it suitable for testing a range of effects of interest, including perceptual effects of scene category[46,47] and cognitive effects related to memory[34,35].

EEG was recorded using a BioSemi ActiveTwo amplifier with 64 Ag/AgCl electrodes positioned according to the International 10–10 system, as well as two mastoid electrodes and horizontal and vertical electro-oculograms. Horizontal and vertical electro-oculograms were recorded from electrodes at the lateral canthi and below the eyes, respectively. The Common Mode Sense and Driven Right Leg electrodes were placed on the left and right sides of electrode POz. EEG was recorded in DC mode at a sampling rate of 1,024 Hz with a 200 Hz low-pass filter and 24-bit AD conversion. EEG was recorded in an acoustically shielded recording booth. Continuous EEG data were only minimally preprocessed prior to sharing: data were re-referenced to the POz electrode, downsampled to 512 Hz, and formatted according to the BIDS standard[48]. Crucially, no attempts were made to improve or modify data quality. The shared dataset thus represents a realistic range of signal quality and typical EEG artifacts.

Additional technical details regarding the behavioral task and EEG dataset, including a detailed explanation of the task structure, experimental conditions, and data organization, were provided to analysts in an accompanying documentation file and are further described in the data descriptor paper[43].

4.3 Hypotheses

Given the task design, the steering committee initially specified eight hypotheses to be tested based on the provided data set[42]. Due to time and resource constraints, these were narrowed to three established ERP hypotheses concerning visual stimulus processing and long-term recognition memory. Hypothesis 1 asked whether different categories of scenes were processed differently at an early time point (N100 component; note that scene categories were not matched for low-level visual features, and hence, such features could likely drive differences in EEG signal across categories). Hypothesis 2 asked which brain processes are involved in detecting image novelty. Hypothesis 3 concerned which EEG features distinguished correct recognition decisions from false alarms. According to their formulation, these hypotheses varied in terms of specificity: Hypotheses 1 and 2 provided a predefined ERP component or range of channels, and time windows to investigate, while hypothesis 3 was more open-ended without mentioning a particular spatial or temporal region of interest.

4.4 Instructions for analyst teams

Analyst teams were instructed to analyze the EEG dataset and test the hypotheses using the same analysis pipelines they would typically apply in their own research. The instructions outlined the project's goals, described the dataset, listed the hypotheses, and specified the submission requirements for results, analysis code, and processed data. The full instructions sent to the analyst teams are available in the data descriptor paper[43] and at https://osf.io/xfrbe/files/4nhck.

Analyst teams were required to submit a comprehensive set of results and documentation. This included: (1) binary decisions and confidence ratings for each hypothesis (prior and posterior questionnaires), (2) a concise free-text report of their results, formatted like a typical results section of a scientific paper, (3) detailed responses to an analysis questionnaire documenting all preprocessing and analysis steps, (4) all analysis scripts needed to reproduce their pipeline, including code documentation or stepwise descriptions for GUI-based analyses, and (5) preprocessed trial-by-trial EEG data for each participant, along with documentation of bad channels, rejected trials, and independent components removed during preprocessing.

4.5 Estimating variability in pipelines and results

It is important to note that we could not fully prevent cases of missing components of reports (e.g., incomplete result reports for each tested hypothesis) or missing parts of the files (e.g., incomplete/partial scripts, empty EEG data within a file, and missing event markers).

4.5.1 Pipeline variability

Based on the scripts provided by the teams, we revised both the individual values and the order of steps in the analysis questionnaire reports and used the corrected values in the variability analyses.

For each analysis step and decision, we quantified decision variability by normalized Shannon entropy, based on the distribution of teams across all decision options. The normalized entropy index is defined as:

$$H_{norm}(X) = \frac{H(X)}{H_{max}} = \frac{-\Sigma_{i=1}^{n} p_i \, log_2(p_i)}{log_2 n},$$

where $H(X)$ is the Shannon entropy of the distribution of each of the processing step choices, n is the total number of categories within the step, and p_i is the proportion of the category i.

This measure approaches 0 when one option is chosen with overwhelming consistency and 1 when all options are chosen equally often (uniform distribution). Because entropy is insensitive to sample size, we supplemented it with exact binomial tests against a 0.5 null. In cases where normalized entropy exceeded 0.98 threshold, we generally observed no statistically significant deviation from uniformity.

To assess the prototypicality of pipelines, we introduced the parameter-choice prototypicality index, which measured how common analysis choices were among teams. For each processing step s, we computed the proportion of teams that made each possible choice. For a given team t, let $p_{t,s}$ denote the proportion of teams that made the same choice as team t at step s. The parameter-choice prototypicality index for team t was then calculated as:

$$CTI_t = \frac{1}{S}\sum_{s=1}^{S} p_{t,s} \ ,$$

where S is the total number of processing steps. Lower values indicate that a team tended to use rarer choices, whereas higher values indicate more common choices.

To quantify the sequence of processing steps, we defined 12 commonly used ERP analysis steps and asked analyst teams to indicate the order in which they performed each one. Steps not used by the team were left blank. Each step corresponded to a specific operation (e.g., high-pass or low-pass filtering, segment rejection, etc.) and the sequence preserved the order of application. Steps used by fewer than 25% of teams were removed (i.e., “Detrending”, “Spatial transform”). To quantify how frequently one preprocessing step followed another, pipelines were encoded as a transition probability matrix: rows correspond to preceding steps, and columns correspond to subsequent steps (see Figure 2A). Each cell value reflects the probability that a given step (column) immediately follows another step (row), calculated as the number of times a given transition occurred across all pipelines. One team did not report the order of steps and one team reported the order of only one step and were consequently excluded from the transition probability matrix.

To estimate the prototypical order of steps, we identified the team whose step sequence had the highest average probability from the transition probability matrix. According to this procedure, the order of the prototypical pipeline was high-pass filtering, low-pass filtering, epoching, baseline correction, and segment exclusion. To compute the distance of the pipelines submitted by all teams with this prototypical order of steps, a letter code was assigned to each preprocessing operation, and the entire pipeline was converted into a string. The Levenshtein string similarity[32] was then calculated by comparing character-by-character each pipeline sequence to the prototypical string: based on the number of insertions, deletions, or substitutions needed to transform one string into another, a distance score from the prototypical pipeline was assigned to each team. The Levenshtein score was used to estimate each teams’ distance to the step-order prototypicality.

4.5.2 EEG time series data

We were unable to process the time-series data from 51 teams for the following reasons: the data were not epoched (n = 23), the files could not be opened (n = 18), or essential information—such as channel labels, time points, data dimensions, or supporting files—was missing (n = 10; a more detailed description can be found in [43]). The EEGManyPipelines steering committee brought data from 117 teams into a matched format: EEG time-series data were transformed to the same epoch window length (-200 to 600 ms), number and order of channels, and sampling rate (256 Hz). If the data had a shorter epoch length or a varying number of participants (due to participant

exclusions), we added NaN values to standardize the format across teams. Thirty-four teams provided their data in volts, which we converted to microvolts.

Difference waves

Out of 117 teams whose data was available, for hypothesis 1 the difference waves could be extracted from 109 teams due to problems with event markers (n = 4), empty data for one of the conditions (n = 2), preprocessing reference to a different hypothesis (n = 1), or problems with channel labels (n = 1). For hypotheses 2 and 3, data from 115 teams were initially available (two datasets from the 117 referred to different hypotheses); a further 27 teams were excluded due to problems with event markers and 3 teams were excluded due to empty data structures, leaving 88 teams whose data could be used for further analyses.

Difference waves were calculated by subtracting the averaged ERPs of the two conditions that are contrasted in each of the three hypotheses. For hypothesis 1, the difference wave was obtained by subtracting the grand-average ERP for natural stimuli from that for man-made stimuli. For hypothesis 2, the grand-average ERP for previously seen ('old') pictures was subtracted from that for new pictures. For hypothesis 3, the grand-average ERP for 'misses' (old pictures falsely identified as new) was subtracted from that for 'hits' (old pictures correctly recognized as old).

Comparing signal variability across teams and participants

To quantify the variability in results induced by different analytic decisions (i.e., across analyst teams) relative to the variability typically observed across the participants in an EEG study (e.g., due to differences in skull size and thickness, brain anatomy, etc.), we computed grand-average ERPs and difference waves for hypothesis 1 across teams and across participants. This was completed to provide an overall representation of the neural responses used to compare EEG signal variability across teams and participants.

To compute variability across participants, we averaged the ERPs and difference waves across analysis teams for each participant and computed the variance of the participant averages. Next, to compute variability across analysis teams, we calculated the grand-average ERPs (average across participants) and difference waves for each analysis team, and computed the variance across analysis teams (see Figure 4). Variance was calculated using an implementation of the *var* function in R that uses the unbiased sample variance estimator, defined as the average squared deviation from the mean with $n - 1$ as denominator, where n is the number of observations.

4.6 The link between prototypical pipelines and the results

To answer whether unusual pipelines were associated with unusual results, we estimated the average difference wave across teams for each hypothesis. For each hypothesis, we computed the median absolute difference wave and averaged it within a 100–600 ms time window. For hypothesis 1 we used channel CPz, and for hypotheses 2 and 3 we used channel Fz. Outliers were identified as mean amplitude below $Q1 - 3 * IQR$ or above $Q3 - 3 * IQR$ and excluded

from further analysis. We then correlated these values with the step-order prototypicality (Levenshtein distance between each team's step order and the prototypical step order) and parameter-choice prototypicality index, using Spearman's Rank correlations.

4.7 Linking analysis choices to results

We used the corrected values from the analysis questionnaire as independent variables in linear models. The dependent variables were derived from the reported *p*-values and difference waves from three hypotheses. Unfortunately, we were unable to use standardized effect sizes as dependent variables due to the high heterogeneity in the reports. Even within the same statistical test (e.g., cluster-based statistics), the measures of the effect can differ, for example, by referring to the range (e.g., *t*-min to *t*-max), the mean (e.g., *t*-mean), or the sum of test statistics (e.g., *t*-sum).

Because the distribution of test statistics was skewed, we transformed *p*-values into z-scores using the inverse normal quantile function in R version 4.4.1[49] to approximate a standard normal distribution. If a team reported multiple significant *p*-values across electrodes and time windows, we took the median *p*-value across their significant findings. In the models that used the transformed *p*-values, model specifications were:

- ***Hypothesis 1.*** Out of 168 teams that completed the analysis, 14 teams did not report a *p*-value, one team did not report if they corrected for multiple comparisons, and one team was removed from the analysis as an outlier identified by studentized residual exceeding an absolute value of 3. We included 152 teams in the linear model. We excluded information on software host, method of multiple comparisons, and high-pass filter direction from linear models due to high collinearity among these variables, as indicated by the variance inflation factor (VIF). For hypothesis 1, independent variables were: software, high-pass cutoff frequency, filter type, reference, sampling frequency, number of excluded subjects, analyzed topographic region, analyzed number of channels, averaging over the channels (a binary outcome, yes/no), segment exclusion criteria, statistical method, correction for multiple comparisons (a binary outcome, yes/no), analyzed time window length and whether it was averaged or not (a binary outcome, yes/no), independent component analysis (ICA) algorithm, bad-component selection visual or plugin based (a binary outcome, yes/no), and baseline window length.

- ***Hypothesis 2.*** Of the 168 teams that completed the analysis, 20 did not report a *p*-value, 23 teams did not report a *p*-value with the necessary precision, and one team did not report whether they averaged across channels. A total of 124 teams provided full reports that were included in the linear model. We excluded information on software host, multiple-comparison method, and high-pass filter direction from linear models due to high collinearity among these variables (based on VIF scores). The independent variables were the same for hypothesis 2 as for hypothesis 1, except for the analyzed time window, as it was already specified in hypothesis 2, and only a small fraction of teams deviated from it.

- ***Hypothesis 3.*** Out of 168 teams that completed the analysis, 52 teams did not report a *p*-value, one team did not report if they averaged across the sensors, one team did not report if they averaged across the time window, and one team was removed from the analysis based on studentized residual exceeding an absolute value of 3. In total, 113 teams submitted full reports and were included in the linear model. We excluded information on software host, software, and high-pass filter direction from linear models due to high collinearity associated with these variables (based on VIF scores). For hypothesis 3, the independent variables were the same as for hypothesis 1, except for software, the number of analyzed channels, and the topographic region. We did not include the last two variables due to missing information.

We used difference waves of the hypothesis-specific conditions as dependent variables in linear models. The model specifications for the three hypotheses were:

- ***Hypothesis 1.*** The mean difference wave from the CPz channel representing the N100 component (80–120 ms) served as the independent variable in the linear models for hypothesis 1. Of the 109 teams with available data, eight lacked the CPz channel and were excluded, and one team was removed as an outlier due to exceptionally high mean amplitude values (above 10 μV). From the remaining 100 teams, four were further excluded based on studentized residuals (exceeding 3). Data from 96 teams were used in the linear model. The independent variables were: software, high-pass cutoff- and type, reference, sampling frequency, number of excluded subjects, segment exclusion criteria, ICA algorithm, visual- or plugin-based bad-component selection (a binary outcome, yes/no), and baseline window length.

- ***Hypothesis 2.*** The mean difference wave from Cz, Fz, FCz, FC1, FC2, FC3, FC4, FC5, and FC6 channels in the time window between 300 and 500 ms was used as the independent variable in linear models for hypothesis 2. Of the 88 teams for which difference waves were available, two teams were excluded based on studentized residuals (exceeding 3), two teams were excluded due to exceptionally high mean amplitude values (above 10 μV), and one team due to extreme leverage, which led to numerical instability in the HC3 robust variance estimation. Data from 83 teams were used in the linear model. The independent variables were: software, high-pass cutoff, type, and direction, reference, sampling frequency, number of excluded subjects, segment exclusion criteria, ICA algorithm, visual- or plugin-based bad-component selection, and baseline window length.

- ***Hypothesis 3.*** The mean difference wave from Fz channels in the time window between 400 and 600 ms was used as the independent variable in the linear model for hypothesis 3. Of the 88 teams for which difference waves were available, two teams were excluded due to exceptionally high mean amplitude values (above 10 μV), one team was excluded based on studentized residuals (exceeding 3), one team was excluded due to extreme leverage, which led to numerical instability in the HC3 robust variance estimation, and one team was excluded due to missing data points. Data from 83 teams were used in the linear model. The independent variables were: software, high-pass cutoff, type, and direction,

reference, sampling frequency, number of excluded subjects, segment exclusion criteria, ICA algorithm, visual- or plugin-based bad-component selection, and baseline window length. To assess the generalizability of the findings, we also tested the mean difference wave in the Pz channel using the same linear model specifications.

For all of linear models, we also included 4 variables representing the order of steps (all teams received a binary score, yes/no, for each category): epoching before high-pass filtering, low-pass filtering before artifact correction/rejection, re-referencing before artifact correction/rejection, and artifact correction before artifact rejection. These categories were chosen based on prior literature[50].

Missing reports from the analysis questionnaire detailing the data processing and analysis pipelines were filled in as a mean score across teams for continuous variables, and as 'unknown' for categorical ones to preserve other pipeline responses in the specified model. Continuous variables were standardised to z-scores (mean = 0, SD = 1).

Because the initial linear model included many predictors, there was a risk of overfitting. To identify the most parsimonious model, stepwise model selection was performed in both directions using the Bayesian Information Criterion (BIC). The stepwise procedure was implemented using the stepAIC() function, with the penalty parameter adjusted as $k = log(n)$ to approximate the BIC for more conservative model selection. The final model was selected using the BIC score, which identifies the model expected to provide the best balance between fit and simplicity. Subsequent coefficient estimates and *p*-values are conditional on this selected model and should be interpreted descriptively rather than as confirmatory tests.

Model assumptions were evaluated prior to interpretation. Influential observations were examined using studentized residuals. Observations with absolute studentized residuals greater than 3 were identified and further assessed for undue influence using leverage and Cook's distance criteria. Linearity was evaluated by examining the relationship between fitted values and residuals; no systematic patterns were observed. Residual normality was assessed using the Shapiro–Wilk and Kolmogorov–Smirnov tests. Homoscedasticity was evaluated using the Breusch–Pagan test. Models that included the difference waves from hypotheses 2 and 3, and the *p*-values from hypothesis 1 as dependent variables, showed evidence of heteroscedasticity. To account for heteroscedasticity, robust standard errors were estimated using the HC3 heteroscedasticity-consistent covariance matrix.

Data availability

The EEGManyPipelines dataset can be downloaded from the brainlife.io platform (https://brainlife.io/project/6863bf5c1521e536327bfea7[51]) under the Creative Commons Attribution (CC-BY) license. The instructions for the analysts detailing the EEG dataset and task, the questionnaires, and forms can also be found at the EEGManyPipelines Open Science Framework repository (https://osf.io/c5hyt). For information on the data structure and usability, we refer the reader to the accompanying data descriptor paper[43].

Code availability

https://github.com/EEGManyPipelines/Main_Paper/tree/main

References

1. Holzmeister, F. *et al.* Heterogeneity in effect size estimates. *Proc. Natl. Acad. Sci. U.S.A.* **121**, e2403490121 (2024). https://doi.org/10.1073/pnas.2403490121

2. Hussey, I. Data is not available upon request. Preprint at https://doi.org/10.31234/osf.io/jbu9r_v2 (2025).

3. Womack, R. P. Research Data in Core Journals in Biology, Chemistry, Mathematics, and Physics. *PLoS ONE* **10**, e0143460 (2015). https://doi.org/10.1371/journal.pone.0143460

4. Yang, P. & Colavizza, G. Research Data in Scientific Publications: A Cross-Field Analysis. Preprint at https://doi.org/10.48550/arXiv.2502.01407 (2025).

5. Styles, S. J., Ković, V., Ke, H. & Šoškić, A. Towards ARTEM-IS: Design guidelines for evidence-based EEG methodology reporting tools. *NeuroImage* **245**, 118721 (2021). https://doi.org/10.1016/j.neuroimage.2021.118721

6. Klein, R. A. *et al.* Investigating Variation in Replicability: A "Many Labs" Replication Project. *Social Psychology* **45**, 142–152 (2014). https://doi.org/10.1027/1864-9335/a000178

7. Klein, R. A. *et al.* Many Labs 2: Investigating Variation in Replicability Across Samples and Settings. *Advances in Methods and Practices in Psychological Science* **1**, 443–490 (2018). https://doi.org/10.1177/2515245918810225

8. Huber, C. *et al.* Competition and moral behavior: A meta-analysis of forty-five crowd-sourced experimental designs. *Proc. Natl. Acad. Sci. U.S.A.* **120**, e2215572120 (2023). https://doi.org/10.1073/pnas.2215572120

9. Landy, J. F. *et al.* Crowdsourcing hypothesis tests: Making transparent how design choices shape research results. *Psychological Bulletin* **146**, 451–479 (2020). https://doi.org/10.1037/bul0000220

10. Lai, C. K. *et al.* Reducing implicit racial preferences: I. A comparative investigation of 17 interventions. *Journal of Experimental Psychology: General* **143**, 1765–1785 (2014). https://doi.org/10.1037/a0036260

11. Milkman, K. L. *et al.* Megastudies improve the impact of applied behavioural science. *Nature* **600**, 478–483 (2021). https://doi.org/10.1038/s41586-021-04128-4

12. Silberzahn, R. *et al.* Many Analysts, One Data Set: Making Transparent How Variations in Analytic Choices Affect Results. *Advances in Methods and Practices in Psychological Science* **1**, 337–356 (2018). https://doi.org/10.1177/2515245917747646

13. Olsson-Collentine, A., Wicherts, J. M. & Van Assen, M. A. L. M. Heterogeneity in direct replications in psychology and its association with effect size. *Psychological Bulletin* **146**, 922–940 (2020). https://doi.org/10.1037/bul0000294

14. Constant, M. *et al.* A multilab investigation into the N2pc as an indicator of attentional selectivity: Direct replication of Eimer (1996). *Cortex* **190**, 304–341 (2025). https://doi.org/10.1016/j.cortex.2025.05.014

15. Botvinik-Nezer, R. *et al.* Variability in the analysis of a single neuroimaging dataset by many teams. *Nature* **582**, 84–88 (2020). https://doi.org/10.1038/s41586-020-2314-9

16. Breznau, N. *et al.* Observing many researchers using the same data and hypothesis reveals a hidden universe of uncertainty. *Proc. Natl. Acad. Sci. U.S.A.* **119**, e2203150119 (2022). https://doi.org/10.1073/pnas.2203150119

17. Menkveld, A. J. *et al.* Nonstandard Errors. *The Journal of Finance* **79**, 2339–2390 (2024). https://doi.org/10.1111/jofi.13337

18. Aczel, B. *et al.* Investigating the analytical robustness of the social and behavioural sciences. *Nature* **652**, 135–142 (2026). https://doi.org/10.1038/s41586-025-09844-9

19. Schweinsberg, M. *et al.* Same data, different conclusions: Radical dispersion in empirical results when independent analysts operationalize and test the same hypothesis. *Organizational Behavior and Human Decision Processes* **165**, 228–249 (2021). https://doi.org/10.1016/j.obhdp.2021.02.003

20. Trübutschek, D. *et al.* EEGManyPipelines: A Large-scale, Grassroot Multi-analyst Study of EEG Analysis Practices in the Wild. *Journal of Cognitive Neuroscience* **36**, 217–224 (2023). https://doi.org/10.1162/jocn_a_02087

21. Luck, S. J. *Applied Event-Related Potential Data Analysis*. (LibreTexts, 2022). https://doi.org/10.18115/D5QG92.

22. Hari, R. & Puce, A. *MEG-EEG Primer*. (Oxford University Press, New York, NY, 2023). https://doi.org/10.1093/med/9780197542187.001.0001.

23. Cohen, M. X. *Analyzing Neural Time Series Data: Theory and Practice*. (The MIT Press, 2014). https://doi.org/10.7551/mitpress/9609.001.0001.

24. Robbins, K. A., Touryan, J., Mullen, T., Kothe, C. & Bigdely-Shamlo, N. How Sensitive Are EEG Results to Preprocessing Methods: A Benchmarking Study. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* **28**, 1081–1090 (2020). https://doi.org/10.1109/TNSRE.2020.2980223

25. Clayson, P. E., Baldwin, S. A., Rocha, H. A. & Larson, M. J. The data-processing multiverse of event-related potentials (ERPs): A roadmap for the optimization and standardization of ERP processing and reduction pipelines. *NeuroImage* **245**, 118712 (2021). https://doi.org/10.1016/j.neuroimage.2021.118712

26. Šoškić, A., Jovanović, V., Styles, S. J., Kappenman, E. S. & Ković, V. How to do Better N400 Studies: Reproducibility, Consistency and Adherence to Research Standards in the Existing Literature. *Neuropsychol Rev* **32**, 577–600 (2022). https://doi.org/10.1007/s11065-021-09513-4

27. Huang, H., Moffa, A. H., Loo, C. & Nikolin, S. No Single Best Pipeline: Multiverse Analysis of EEG Preprocessing for N-Back Working Memory Tasks. *Psychophysiology* **62**, e70197 (2025). https://doi.org/10.1111/psyp.70197

28. Cyrus-Lai, W., Tierney, W., Schweinsberg, M. & Uhlmann, E. L. Exposing and overcoming the fixed-effect fallacy through crowd science. *Behav Brain Sci* **45**, e8 (2022). https://doi.org/10.1017/S0140525X21000297

29. Auspurg, K. & Brüderl, J. Has the Credibility of the Social Sciences Been Credibly Destroyed? Reanalyzing the “Many Analysts, One Data Set” Project. *Socius: Sociological Research for a Dynamic World* **7**, 23780231211024421 (2021). https://doi.org/10.1177/23780231211024421

30. Auspurg, K. & Brüderl, J. Toward a more credible assessment of the credibility of science by many-analyst studies. *Proc. Natl. Acad. Sci. U.S.A.* **121**, e2404035121 (2024). https://doi.org/10.1073/pnas.2404035121

31. Paul, M. & Mani, N. Preprocessing and analysis practices in developmental N400 research – a systematic review and pipeline comparison. Preprint at https://doi.org/10.31234/osf.io/j235p (2022).

32. Fahimi, R. & Bruce, N. D. B. On metrics for measuring scanpath similarity. *Behav Res* **53**, 609–628 (2021). https://doi.org/10.3758/s13428-020-01441-0

33. Pion-Tonachini, L., Kreutz-Delgado, K. & Makeig, S. ICLabel: An automated electroencephalographic independent component classifier, dataset, and website. *NeuroImage* **198**, 181–197 (2019). https://doi.org/10.1016/j.neuroimage.2019.05.026

34. Voss, J. L. & Paller, K. A. Neural Substrates of Remembering: Event-Related Potential Studies ☆. in *Learning and Memory: A Comprehensive Reference* 81–98 (Elsevier, 2017). https://doi.org/10.1016/B978-0-12-809324-5.21070-5.

35. Wilding, E. L. & Ranganath, C. *Electrophysiological Correlates of Episodic Memory Processes*. (Oxford University Press, 2011). https://doi.org/10.1093/oxfordhb/9780195374148.013.0187.

36. Jovanović, V., Petrušić, I., Ković, V. & Savić, A. M. The Practical Implications of Re-Referencing in ERP Studies: The Case of N400 in the Picture–Word Verification Task. *Diagnostics* **15**, 156 (2025). https://doi.org/10.3390/diagnostics15020156

37. Šoškić, A., Styles, S. J., Kappenman, E. S. & Ković, V. Garden of forking paths in ERP research – Effects of varying pre-processing and analysis steps in an N400 experiment. *Psychophysiology* **61**, e14628 (2024). https://doi.org/10.1111/psyp.14628

38. Acunzo, D. J., MacKenzie, G. & Van Rossum, M. C. W. Systematic biases in early ERP and ERF components as a result of high-pass filtering. *Journal of Neuroscience Methods* **209**, 212–218 (2012). https://doi.org/10.1016/j.jneumeth.2012.06.011

39. Tanner, D., Morgan-Short, K. & Luck, S. J. How inappropriate high-pass filters can produce artifactual effects and incorrect conclusions in ERP studies of language and cognition: High-pass filtering and artifactual ERP effects. *Psychophysiology* **52**, 997–1009 (2015). https://doi.org/10.1111/psyp.12437

40. Zhang, G., Garrett, D. R. & Luck, S. J. Optimal filters for ERP research II : Recommended settings for seven common ERP components. *Psychophysiology* **61**, e14530 (2024). https://doi.org/10.1111/psyp.14530

41. Šoškić, A. *et al. ARTEM-IS for ERP: Agreed Reporting Template for EEG Methodology - International Standard for Documenting Studies on Event-Related Potentials*. Preprint at https://doi.org/10.31234/osf.io/mq5sy (2023).

42. Algermissen, J. *et al.* EEGManyPipelines: Robustness of EEG results across analysis pipelines. Preprint at https://doi.org/10.17605/OSF.IO/42K5H (2022).

43. Cesnaite, E. *et al.* The EEGManyPipelines Dataset: Metascientific Data on 168 Independent Analyses of a Single EEG Dataset. Preprint at https://doi.org/10.31222/osf.io/c4xwg_v1 (2025).

44. Shepard, R. N. & Teghtsoonian, M. Retention of information under conditions approaching a steady state. *Journal of Experimental Psychology* **62**, 302–309 (1961). https://doi.org/10.1037/h0048606

45. Friedman, D. Cognitive Event-Related Potential Components During Continuous Recognition Memory for Pictures. *Psychophysiology* **27**, 136–148 (1990). https://doi.org/10.1111/j.1469-8986.1990.tb00365.x

46. Groen, I. I. A., Ghebreab, S., Prins, H., Lamme, V. A. F. & Scholte, H. S. From Image Statistics to Scene Gist: Evoked Neural Activity Reveals Transition from Low-Level Natural Image Structure to Scene Category. *J. Neurosci.* **33**, 18814–18824 (2013). https://doi.org/10.1523/JNEUROSCI.3128-13.2013

47. Hansen, B. C., Johnson, A. P. & Ellemberg, D. Different spatial frequency bands selectively signal for natural image statistics in the early visual system. *Journal of Neurophysiology* **108**, 2160–2172 (2012). https://doi.org/10.1152/jn.00288.2012

48. Pernet, C. R. *et al.* EEG-BIDS, an extension to the brain imaging data structure for electroencephalography. *Sci Data* **6**, 103 (2019). https://doi.org/10.1038/s41597-019-0104-8

49. R Core Team. R: A Language and Environment for Statistical Computing. (R Foundation for Statistical Computing, Vienna, Austria, 2024).

50. Luck, S. J. *An Introduction to the Event-Related Potential Technique*. (The MIT Press, Cambridge, Massachusetts, 2014).

51. Hayashi, S. *et al.* brainlife.io: a decentralized and open-source cloud platform to support neuroscience research. *Nat Methods* **21**, 809–813 (2024). https://doi.org/10.1038/s41592-024-02237-2

List of authors

Elena Cesnaite[1], Johannes Algermissen[2,3,4], Mikkel C. Vinding[5,6], Andrea Vitale[7], Annalisa Pascarella[8], Nastassja L. Fischer[9,10], Yu-Fang Yang[11], Darinka Trübutschek[12,13], Claudia Gianelli[14], Charlotte R. Marshall[15], Omid Abbasi[16], Timofey Adamovich[17], Lena Adel[18], Nikita Agarwal[19,20], Blanca Aguado-López[21], Uma Ajmeria[22], Esra Al[23], Hannah Aldeen[24], Ricardo J. Alejandro[25], Evgeniia A. Alenina[26], Meera Alex[27], Ettore Ambrosini[28], Douglas J. Angus[29], Alessandra Anzolin[30], Stefan Appelhoff[31], Giorgio Arcara[32,33], Yana Arkhipova[34], Maria Azanova[23,35], Chiara Bagattini[36], Daniel H. Baker[37], Christoph Bamberg[38], Francesca M. Barbero[39], Ceren Battal[40,39], Anna-Katharina R. Bauer[41], Burcu Bayram[42], Paulo R. Bazán[43], Henry Beale[44], Elena Belanova[45], Zara M. Bergström[46], Giulio Bernardi[47], Martina Berto[47], Monica Betta[47], Isabelle Blanchette[48], Marta Bortoletto[47], Davide Bottari[47], Claire Braboszcz[49], Daniel E. Bradford[50], Joseph L. Brooks[51], Chloe M. Brunskill[52], Florian Bublatzky[53,54], Daniel Büchel[55], Roberta P. Calce[56,39], Mariagrazia Capizzi[21,57], Raymundo Cassani[58], María Concepción Castellanos[21], Luigi Cattaneo[59], Xim Cerda-Company[60,61], Lixiang Chen[62], Xinyuan Chen[63], Yiyu Chen[64], Yoojeong Choo[65,66], Hu Chuan-Peng[67], Karan Chugani[49], Radoslaw Martin Cichy[68], Barbora Cimrová[69], Simon Ciranka[31], Oren Civier[70], Carlotta Cogoni[71], Michel-Pierre Coll[48], Julie Coloigner[72], Angela Conejero[21], Martin Constant[73,74], César E. Corona-González[75,76], Antonio Criscuolo[77], Damian Cruse[78], Claire Cury[72], Artur Czeszumski[79], Martin J. Dahl[80], Matthew Davidson[81], Ingmar E. J. de Vries[59,3], M. Dolores del Castillo[82], Vincent DeLuca[83], Gianpaolo Demarchi[84], Kobe Desender[85], Marcelo Dias[86], Darcy Waller[87], Martin Dietz[88], Olaf Dimigen[89], Hao Ding[90,91], Li Dong[92], Linda Drijvers[3,93], Brandi Lee Drisdelle[94,95], Gian Marco Duma[96], Guillaume Dumas[97], Benedikt Ehinger[98], Alexandra K. Emmendorfer[99], Alexander Enge[100], Stefanie Evas[101,102], Tommaso Fedele[103], Alessandra Federici[47], Gordon B. Feld[104], Lauren K. Fink[105], Shai Fischer[106], Emilia Fló[107,108], Silvia Formica[109], Norman Forschack[110], Bettina Forster[111], Marcel Franz[112], Léon Franzen[113], Cassidy M. Fry[114], Alejandro Galvez-Pol[115], Haydee G. García-Lázaro[116,117], José Carlos García Alanis[118], Patricia Garrido-Vásquez[119], Billy Gerdfeldter[120], Georgia Gerike[121,122], Magdalena Gippert[23], Carlos González-García[21], Dario Gordillo[123], Anna Grabowska[124,125], Lisa-Marie Greenwood[126], Benjamin J. Griffiths[22], Demetrio Grollero[127], Joachim Gross[16], Danièle A. Gubler[128], Pierre Guilleminot[129], Berna Güler[130], Christopher Gundlach[110], Eren Günseli[130], Mingqian Guo[131], Malte R. Güth[132], Céline Haciahmet[133], Itay Hadas[134], Jarmo Hämäläinen[121], Sizhu Han[135], Anthony M. Harris[44], Mahmoud Hassan[136,137], Stefan Haufe[138,19,139], Anibal S. Heinsfeld[140], Peer Herholz[141], Mario Hervault[142], Aron T. Hill[143], Alice Hodapp[34], Maximilian Hommelsen[144], Hanna Honcamp[145], Thomas J. Hosang[146], Chun-Hsien Hsu[147], Tzu-Yu Hsu[148,147], Changrun Huang[149], Christoph Huber-Huber[59], Matthew Hughes[70], Agustin Ibanez[150,151,152,153,154], Behzad Iravani[155], Kylie Isenburg[30], Bradley N. Jack[126], Mina Jamshidi Idaji[23,156], Oskar H. Jefsen[157], El Jeong[158], Fang Jiang[159], Ann-Kathrin Joechner[80], Matthew R. Johnson[160], Ki-Young Jung[161], Neda Kaboodvand[155], Daniel Kaiser[162,163,164,165], Patrycja Kałamała[125], Taeho Kang[166], Nikolai Kapralov[167,168], Hamid Karimi-Rouzbahani[169,170], Mikolaj Kegler[171], Nicholas J. Kelley[172], Simon Kern[104], Casper Kerrén[23], Songhee Kim[173], Alina Kiseleva[103], Laura-Isabelle Klatt[174], Felix Klotzsche[23], Daniel S. Kluger[175], Peter König[176,177], Bruno Kopp[178], Julian Q. Kosciessa[3,31], Maciej Kosilo[71], Boris Kotchoubey[179], Layla Kouara[180], Sander Krewinkel, Leon O. H. Kroczek[181], Corinna Kührt[182,183], Sara Lago[32,184], Martin Lamoš[185], Nicolas Langer[186], Dmitry Lazurenko[187], Pierre Le Denmat[155], Ty Lees[188,189], Tim Lehmann[55], Tomas Lenc[39,190], Cédric Lenoir[39], Wanru Li[191], Zefeng Li[192], Anna M. Liesefeld[193], Heinrich R. Liesefeld[194], Maxim Likhanov[195], Evan N. Lintz[196], Giulia Lioi[197], Vladimir Litvak[198], Xiaoyi Liu[199], Yikang Liu[67], Yu-Hui Lo[200], Eduardo López-Caneda[201,202], David López-García[21], Corinna Lorenz[203], Andreas Löw[146], Runhao Lu[169,204], Zitong Lu[205], Alicia Luque[206], David Lydon-Staley[207], Oskar MacGregor[208], Antonio Maffei[209,210], Martin E. Maier[211], Lisa M. Makowski[212], Francesco Mantegna[213,214], Eleonora Marcantoni[215], María Carmen Martín-Buro[216], Ana Martinovici[217], Anna

Marzecova[218], Fabio Masina[33], Jason B. Mattingley[44], Sara Mazzini[93], Takfarinas Medani[219], Ting Mei[131], David Melcher[220], Qingqing Meng[221], Nick S. Menger[222], Ahmad Mheich[136,223], Milan Mitka[69], Claire Monroy[51], Maria Montefinese[224], David Moreau[225], Quentin Moreau[226], Liad Mudrik[227], Faisal Mushtaq[180], Nicholas E. Myers[22], Norberto E. Naal-Ruiz[75], Foroogh Najafi[228], Hong-Viet V. Ngo[229], Lu Nie[230], Martin Nilsson[231], Guiomar Niso[232,233], Erika Nyhus[234], Joyce Oerlemans[25], Emanuele Olivetti[235], Joan Orpella[213,236], Eduard Ort[218], Ana F. Palenciano[21], Kyoungeun Park[237], Bernhard Pastötter[238], Zita Patai[239,240], Katharina Paul[241], Artur José M. Paulo[43], Giovanni Pellegrino[242], Sergio M. Pereira Soares[93], Franco Pestilli[243], Rachele Pezzetta[32], Daniela M. Pfabigan[244], Anna Plachti[245], Ulrich Pomper[246], Tzvetan Popov[186], Pavel Prado[247], Andrea B. Protzner[248], Yanina Prystauka[249], Haoyue Qian[250], Milena Rabovsky[34], Janir Ramos da Cruz[251,252], Manuel Rausch[253,254,255], Amit Rawal[256,257], Stefan Repplinger[258], Danielle R. Rice[259], Igor Riecansky[69], Pavel Riha[185], Johannes Rodrigues[260], M. Paula Roncaglia[261], Michael Rose[262], Marlene Rösner[263,264], Jason Rothman[265,266,267], Simon Ruch[268], Philipp Ruhnau[269], Adi Sarig Golan[270], Arkaprovo Sarkar[271], Akul Satish[169], Cristina Scarpazza[33,32], Christoph Scheffel[182], Judith Schepers[272], Antonio Schettino[273,274], Barbara Schmidt[275], Fabian Schmidt[38], Anna-Lena Schubert[118], Anna-Lisa Schuler[32,23], Catriona L. Scrivener[276], Magdalena Senderecka[125], Ulrike Senftleben[277], Fatih Serin[169], J. Ignacio Serrano[82], Mehran Shabanpour[278], Veronika Shamova[139], Idris Shareef[159], Dmitry Sherbina[187], Cassie Ann Short[279], Carine Signoret[280], Esaú V. P. Sirius[281,43], René Skukies[282], Aureli Soria-Frisch[49], Firat Soylu[283], Sebastian Speer[284], Tim Paul Steinfath[23,168], Robert Steinhauser[211], Simon R. Steinkamp[245], Tilman Stephani[3,23], Alexander Strobel[182], Dawid Strzelczyk[186], Pheobe Sun[285], Yanan Sun[286], Caroline Surrey[182], Nikolay Syrov[287,288], Marco Tagliaferri[59], Siddharth Talwar[289], Vincenza Tarantino[290], Alessandro Tavano[291,292], Jason R. Taylor[293], Carsten Thiele[294], Nivethida Thirugnanasambandam[295], Lara Todorova[130], Aleksandra Tomić[83], Stefan J. Troche[212], Lin-Yuan Tseng[296], Sarah Tune[297], Gözem Turan[298], Luca Turella[59], José Luis Ulloa[299], Michael Valiadis[300], Antonino Vallesi[301], Gwen van der Wijk[248], Marijn van Vliet[302], Marijn van Wingerden[261], Enrico Varano[171], Mohith M. Varma[169,303], Margarida Vasconcelos[304], Yoana Vergilova[305], Adrià Vilà-Balló[306], Antonino Visalli[307], Giada Viviani[210], Robin Vloeberghs[85], Jan Wacker[308], Ya-Jie Wang[309], Meng-Yun Wang[310,3], Tianlu Wang[311], Duan Wei[312], Markus Werkle-Bergner[80], David White[313], Simon Whitton[314], Kristina Wiebels[225], Stefan Wiens[120], Christoph A. Wittkamp[262], Maren-Isabel Wolf[262], Audrey Wong-Kee-You[116], Will Woods[70], Franz Wurm[315,316], Syanah C. Wynn[234], Jiushu Xie[67], Alba Xifra-Porxas[45], Weiyong Xu[121], Lev Yakovlev[317], Jinbiao Yang[93], Mustafa Yavuz[318], Vahab Youssofzadeh[173], Rongjun Yu[319], Talha Zafar[248], Sara Zago[32], Ilya Zakharov[320], Marta Z. Zakrzewska[321], Juanli Zhang[322,23], Zhen Zhang[131], Shanshan Zhen[323], Xinqi Zhou[324], Judy D. Zhu[325], Siyu Zhu[326], Qian Zhuang[327], Catharina Zich[328], Reza Zomorrodi[329], Camila Zugarramurdi[107], Muhammad S. Navid[330,331], Tuomas Puoliväli[332], Mehdi Senoussi[25,333,25], Jeremy D. Yeaton[334], Balazs Aczel[335], Mike X. Cohen[336], Arnaud Delorme[337,338], Anna Dreber[339], Jörg Hipp[340], Felix Holzmeister[341], Magnus Johannesson[339], Vanja Ković[342], Robert Oostenveld[3,6], Yuri G. Pavlov[343], Cyril Pernet[344], Russell A. Poldrack[345], Aina Puce[346], Tom Schonberg[347,270], Martin Schweinsberg[348], Anđela Šoškić[342], Barnabas Szaszi[349,350], Eric L. Uhlmann[351], Gustav Nilsonne[155,120] *, Niko A. Busch[1] *

*shared last authorship

Correspondence and requests for materials should be addressed to Gustav Nilsonne or Niko A. Busch.

[1]Institute of Psychology, University of Muenster, Muenster, Germany.

[2]Department of Experimental Psychology, University of Oxford, Oxford, United Kingdom.

[3]Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen, The Netherlands.

[4]Zurich Center for Neuroeconomics, Department of Economics, University of Zürich, Zurich, Switzerland.

[5]Department of Psychology, University of Copenhagen, Copenhagen, Denmark.

[6]NatMEG, Department of Clinical Neuroscience, Karolinska Institutet, Stockholm, Sweden.

[7]Department of Psychology and Cognitive Science, University of Trento, Italy.

[8]Institute for Applied Mathematics "Mauro Picone" (IAC), CNR, Italy.

[9]Centre for Research in Pedagogy and Practice (CRPP), National Institute of Education (NIE), Singapore.

[10]Science of Learning Centre (SoLEC), National Institute of Education (NIE), Singapore.

[11]Division of Experimental Psychology and Neuropsychology, Department of Education and Psychology, Freie Universität Berlin, Berlin, Germany.

[12]Department of Cognitive Neuroscience, Faculty of Psychology and Neuroscience, Maastricht University, Maastricht, The Netherlands.

[13]Research Group Neural Circuits, Consciousness and Cognition, Max Planck Institute for Empirical Aesthetics, Frankfurt/Main, Germany.

[14]Department of Clinical and Experimental Medicine, University of Messina, Italy.

[15]Centre for Human Brain Health, School of Psychology, University of Birmingham, Birmingham, United Kingdom.

[16]Institute for Biomagnetism and Biosignal Analysis, University of Muenster, Germany.

[17]Federal Scientific Center for Psychological and Interdisciplinary Research, Russian Federation.

[18]McGill University, Canada.

[19]Physikalisch-Technische Bundesanstalt, Germany.

[20]Technische Universität Berlin, Berlin, Germany.

[21]Mind, Brain and Behavior Research Center (CIMCYC), University of Granada, Granada, Spain.

[22]University of Nottingham, United Kingdom.

[23]Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany.

[24]University of Nebraska - Lincoln, United States.

[25]Department of Experimental Psychology, Ghent University, Ghent, Belgium.

[26]Affective Psychophysiology Laboratory, Institute of Health Psychology, HSE University, Saint Petersburg, Russian Federation.

[27]American University of Sharjah, Sharjah, United Arab Emirates.

[28]Department of Neuroscience, University of Padova, Italy.

[29]School of Psychology, Bond University, Gold Coast, Australia.

[30]Department of Physical Medicine and Rehabilitation, Spaulding Rehabilitation Hospital, Harvard Medical School, Boston, MA, United States.

[31]Max Planck Institute for Human Development, Berlin, Germany.

[32]IRCCS San Camillo Hospital, Venice, Italy.

[33]Department of General Psychology, University of Padova, Italy.

[34]Department of Psychology, University of Potsdam, Germany.

[35]Max Planck School of Cognition, Leipzig, Germany.

[36]Department of Neuroscience, Biomedicine and Movement Sciences, University of Verona, Verona, Italy.

[37]Department of Psychology, University of York, United Kingdom.

[38]Center for Cognitive Neuroscience, University of Salzburg, Salzburg, Austria.

[39]Institute of Neuroscience (IoNS), University Catholic of Louvain (UCLouvain), Brussels, Belgium.

[40]Faculty of Psychology and Education Sciences, University of Coimbra, Coimbra, Portugal.

[41]Department of Psychology, Royal Holloway University of London, Egham, United Kingdom.

[42]Department of Cognition, Emotion, and Methods in Psychology Faculty of Psychology University of Vienna Liebiggasse 5 1010 Vienna, Austria.

[43]Hospital Israelita Albert Einstein, São Paulo, Brazil.

[44]Queensland Brain Institute & School of Psychology, The University of Queensland, St Lucia, Australia.

[45]School of Human Sciences, University of Greenwich, London, United Kingdom.

[46]School of Psychology, University of Kent, United Kingdom.

[47]MoMiLab, IMT School for Advanced Studies Lucca, Lucca, Italy.

[48]School of Psychology, Laval University, Quebec City, Canada.

[49]Starlab Barcelona S.L., Spain.

[50]School of Psychological Science, Oregon State University, Corvallis, Oregon, United States.

[51]School of Psychology, Keele University, Keele, United Kingdom.

[52]University of York, United Kingdom.

[53]Department of Psychosomatic Medicine and Psychotherapy, Central Institute of Mental Health Mannheim, Medical Faculty Mannheim/Heidelberg University, Germany.

[54]German Center for Mental Health, partner site Mannheim-Heidelberg-Ulm, Germany.

[55]Exercise Science & Neuroscience, Department Exercise & Health, Paderborn University, Germany.

[56]Institute of Research in Psychology (IPSY), University Catholic of Louvain (UCLouvain), Brussels, Belgium.

[57]Department of Experimental Psychology, University of Granada, Granada, Spain.

[58]Montreal Neurological Institute, McConnell Brain Imaging Centre, McGill University, Montreal, Quebec, Canada.

[59]Center for Mind/Brain Sciences (CIMeC), University of Trento, Italy.

[60]Computer Science Department, Universitat Autonoma de Barcelona, Cerdanyola del Valles, Spain.

[61]Bridging Research in AI and Neuroscience, Computer Vision Center, Cerdanyola del Valles, Spain.

[62]Freie Universität Berlin, Berlin, Germany.

[63]Donders Institute for Brain, Cognition, and Behaviour, Radboud University, Nijmegen, the Netherlands.

[64]Korea University, Republic of Korea.
[65]University of Iowa, Iowa City, IA, United States.
[66]University of Maryland, College Park, MD, United States.
[67]School of Psychology, Nanjing Normal University, China.
[68]Department of Education and Psychology, Freie Universität Berlin, Berlin, Germany.
[69]Centre of Experimental Medicine, Slovak Academy of Sciences, Slovak Republic.
[70]Swinburne University of Technology, Melbourne, Australia.
[71]Instituto de Biofísica e Engenharia Biomédica, Faculdade de Ciências da Universidade de Lisboa, Portugal.
[72]University of Rennes, CNRS, Inria, Inserm, France.
[73]University of Bremen, Germany.
[74]Faculty of Psychology and Educational Sciences, University of Geneva, Geneva, Switzerland.
[75]Tecnologico de Monterrey, School of Humanities and Education, Monterrey, N.L., México.
[76]Department of Physics and Mathematics, Universidad de Monterrey, Nuevo León, Mexico.
[77]Faculty of Psychology and Neuroscience, Maastricht University, Maastricht, The Netherlands.
[78]School of Psychology, University of Birmingham, Birmingham, United Kingdom.
[79]Faculty of Philosophy, University of Warsaw, Warsaw, Poland.
[80]Center for Lifespan Psychology, Max Planck Institute for Human Development, Berlin, Germany.
[81]Discipline of Psychology, University of Technology Sydney, Sydney, Australia.
[82]Center for Automation and Robotics (CAR), Spanish National Research Council (CSIC), Spain.
[83]UiT The Arctic University of Norway, Norway.
[84]Center for Cognitive Neuroscience and Department of Psychology, University of Salzburg, Salzburg, Austria.
[85]Brain & Cognition, KU Leuven, Belgium.
[86]Behavior and Metabolism Laboratory, Champalimaud Research, Champalimaud Foundation, Lisbon, Portugal.
[87]Department of Neuroscience, Brown University, Providence, RI, United States.
[88]Centre for Functionally Integrative Neuroscience, Institute of Clinical Medicine, Aarhus University, Denmark.
[89]University of Groningen, The Netherlands.
[90]Academic Unit of Neurology, Trinity College Dublin, Dublin, Ireland.
[91]Department of Neurology, Universitätsklinikum Würzburg, Würzburg, Germany.
[92]University of Electronic Science and Technology of China, School of Life Science and Technology, Chengdu, China.
[93]Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands.
[94]University of Bristol, Bristol, United Kingdom.
[95]Birkbeck, University of London, London, United Kingdom.
[96]Scientific Institute IRCCS E.Medea, Conegliano (TV), Italy.

[97]CHU Sainte-Justine Azrieli Research Center, Department of Psychiatry and Addictology, University of Montréal, Montréal, Québec, Canada.

[98]University of Stuttgart, Germany.

[99]Section Teaching and Innovation of Learning, Faculty of Psychology and Neuroscience, Maastricht University, Maastricht, The Netherlands.

[100]Department of Psychology, Humboldt-Universität zu Berlin.

[101]School of Psychology, University of Adelaide, Adelaide, Australia.

[102]Human Health, Commonwealth Scientific and Industrial Research Organisation (CSIRO), Adelaide, Australia.

[103]Institute of Cognitive Neuroscience, Higher School of Economics, Russian Federation.

[104]Department of Clinical Psychology, Central Institute of Mental Health Mannheim, Medical Faculty Mannheim/Heidelberg University, Germany.

[105]Department of Psychology, Neuroscience & Behaviour, McMaster University, Hamilton, ON, Canada.

[106]School of Psychological Sciences, Tel Aviv University, Tel Aviv, Israel.

[107]Facultad de Psicología, Universidad de la República, Montevideo, Uruguay.

[108]Institut du Cerveau et de la Moelle Épinière, Paris, France.

[109]Department of Psychology, Humboldt-Universität zu Berlin, Berlin, Germany.

[110]Wilhelm Wundt Institute for Psychology, Leipzig University, Germany.

[111]City St George's, University of London, Centre for Clinical, Social and Cognitive Neuroscience Research, London, United Kingdom.

[112]Department of Clinical Psychology, Friedrich Schiller University Jena, Jena, Germany.

[113]Department of Psychiatry and Psychotherapy, University of Lübeck, Lübeck, Germany.

[114]Department of Human Development and Family Studies, Pennsylvania State University, University Park, Pennsylvania, United States.

[115]University of the Balearic Islands, Psychology Department. Active Cognition, Embodiment, and Environment Lab, Spain.

[116]Smith-Kettlewell Eye Research Institute, United States.

[117]Cardiff University Brain Research Imaging Centre (CUBRIC), School of Psychology, Cardiff University, Cardiff, United Kingdom.

[118]Johannes Gutenberg University Mainz, Germany.

[119]Department of Psychology, University of Concepción, Concepción, Chile.

[120]Department of Psychology, Stockholm University, Stockholm, Sweden.

[121]Department of Psychology, University of Jyväskylä, Jyväskylä, Finland.

[122]Niilo Mäki Institute, Jyväskylä, Finland.

[123]Brain Mind Institute, EPFL, Lausanne, Switzerland.

[124]Doctoral School in the Social Sciences, Jagiellonian University, Cracow, Poland.

[125]Centre for Cognitive Science, Jagiellonian University, Cracow, Poland.

[126]Research School of Psychology, Australian National University, Canberra, Australia.

[127]IMT School for Advanced Studies Lucca, Lucca, Italy.

[128]University of Bern, Switzerland.

[129]Institut de Neuroscience des Systèmes, INSERM, Marseille, France.

[130]Department of Psychology, Sabancı University, Istanbul, Turkey.

[131]Radboud University, The Netherlands.

[132]Center for Molecular and Behavioral Neuroscience, Rutgers University, Newark, NJ, United States.

[133]Department of Cognitive Psychology, University Trier, Germany.

[134]University of California San Diego, La Jolla, California, United States.

[135]Department of Psychology, Philipps-Universität Marburg, Marburg, Germany.

[136]MINDIG, F-35000, Rennes, France.

[137]School of Science and Engineering, Reykjavik University, Reykjavik, Iceland.

[138]Uncertainty, Inverse Modeling and Machine Learning Group, Technische Universität Berlin, Berlin, Germany.

[139]Department of Neurology, Charité Universitätsmedizin Berlin, Berlin, Germany.

[140]Department of Psychology, University of Texas at Austin, Texas, United States.

[141]NeuroDataScience - ORIGAMI laboratory, McConnell Brain Imaging Centre, The Neuro (Montreal Neurological Institute-Hospital), Faculty of Medicine and Health Sciences, McGill University, Montreal, Quebec, Canada.

[142]Univ. Grenoble Alpes, Inserm U1216, CHU Grenoble Alpes, Grenoble Institut Neurosciences, France.

[143]School of Psychology, Deakin University, Melbourne, Australia.

[144]Cognitive Neuroscience, Institute of Neuroscience and Medicine (INM-3), Forschungszentrum Jülich, Jülich, Germany.

[145]Department of Neuropsychology and Psychopharmacology, Faculty of Psychology and Neuroscience, Maastricht University, Maastricht, The Netherlands.

[146]Experimental Psychology Unit, Helmut Schmidt University, Hamburg, Germany.

[147]Institute of Cognitive Neuroscience, National Central University, Taiwan.

[148]Graduate Institute of Mind, Brain and Consciousness, Taipei Medical University, Taipei, Taiwan.

[149]Department of Psychology & Neuroscience, Duke University, Durham, NC, United States.

[150]Latin American Brain Health Institute (BrainLat), Universidad Adolfo Ibañez, Santiago de Chile, Chile.

[151]Global Brain Health Institute, Trinity College Dublin, Dublin, Ireland.

[152]Department of Biophysics, School of Medicine, Istanbul Medipol University, Istanbul, Turkey.

[153]Barcelonaβeta Brain Research Center (BBRC), Pasqual Maragall Foundation, Barcelona, Spain.

[154]Cognitive Neuroscience Center (CNC), Universidad de San Andrés, Buenos Aires, Argentina.

[155]Department of Clinical Neuroscience, Karolinska Institutet, Stockholm, Sweden.

[156]The Berlin Institute for the Foundations of Learning and Data, Berlin, Germany.

[157]Psychosis Research Unit, Aarhus University Hospital – Psychiatry, Denmark.

[158]Seoul National University, Republic of Korea.

[159]Department of Psychology, University of Nevada, Reno, United States.

[160]Inciteful, LLC, United States.

[161]Seoul National University College of Medicine, Republic of Korea..

[162]Neural Computation Group, Department of Mathematics and Computer Science, Physics, Geography, Justus Liebig University Giessen, Germany.

[163]Center for Mind, Brain and Behavior, Universities of Giessen, Marburg, and Darmstadt, Germany.

[164]Cluster of Excellence "The Adaptive Mind", Universities of Giessen, Marburg, and Darmstadt, Germany.

[165]Center for Applied Computer Science and Data Science, Justus Liebig University Giessen, Germany.

[166]Human-Machine Interaction Group, TU Wien, Vienna, Austria.

[167]Department of Neurology, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany.

[168]International Max Planck Research School NeuroCom, Leipzig, Germany.

[169]MRC Cognition and Brain Sciences Unit, University of Cambridge, Cambridge, United Kingdom.

[170]Queensland Brain Institute, University of Queensland, Brisbane, Australia.

[171]Department of Bioengineering, Imperial College London, London, United Kingdom.

[172]University of Southampton, United Kingdom.

[173]Neurology department, Medical College of Wisconsin, Milwaukee, WI, United States.

[174]Leibniz Research Centre for Working Environment and Human Factors, Dortmund, Germany.

[175]Institute for Biomagnetism and Biosignal Analysis, University of Muenster, Muenster, Germany.

[176]University Osnabrück, Germany.

[177]University Medical Center Hamburg-Eppendorf, Hamburg, Germany.

[178]Hannover Medical School, Germany.

[179]Institute of Medical Psychology and Behavioral Neurobiology, University of Tübingen, Tübingen, Germany.

[180]School of Psychology, University of Leeds, Leeds, United Kingdom.

[181]Department of Psychology, Clinical Psychology and Psychotherapy, University of Regensburg, Germany.

[182]Faculty of Psychology, TUD Dresden University of Technology, Germany.

[183]Centre for Tactile Internet with Human-in-the-Loop (CeTI), TUD Dresden University of Technology, Germany.

[184]University of Padova, Italy.

[185]Brain and Mind Research Program, Central European Institute of Technology, Masaryk University, Brno, Czechia.

[186]University of Zurich, Switzerland.

[187]Research Centre for Neurotechnology, Southern Federal University, Rostov-on-Don, Russian Federation.

[188]Center for Depression, Anxiety, and Stress Research, Mclean Hospital, Belmont, United States.

[189]Department of Psychiatry, Harvard Medical School, Boston, MA, United States.

[190]Basque Center on Cognition, Brain and Language (BCBL), Donostia-San Sebastian, Spain.

[191]Peking-Tsinghua Center for Life Sciences, Peking University, Beijing, China.

[192]Ghent University, Belgium.

[193]Department of Psychology, Ludwig-Maximilians-Universität München, Germany.

[194]Department of Psychology, University of Bremen, Germany.

[195]Centre de Recherche en Psychologie et Neurosciences, Aix Marseille University, CNRS, CRPN, Marseille, France.

[196]Sonoma State University, Rohnert Park, CA, United States.

[197]IMT Atlantique, Lab-STICC UMR CNRS 6285, Brest, France.

[198]UCL Queen Square Institute of Neurology, United Kingdom.

[199]Princeton University, United States.

[200]Department of Psychology, National Taiwan University, Taiwan.

[201]Psychology Research Center (CIPsi), University of Minho, Braga, Portugal.

[202]Department of Clinical Psychology and Psychobiology, Faculty of Psychology, Health Research Institute of Santiago de Compostela (IDIS), University of Santiago de Compostela, Spain.

[203]Department of Psychology, University of Wuppertal, Wuppertal, Germany.

[204]Montreal Neurological Institute, Department of Neurology and Neurosurgery, McGill University, Montreal, Quebec, Canada.

[205]MIT McGovern Institute for Brain Research, Boston, United States.

[206]Department of Applied Language Studies and Nebrija Research Center in Cognition, Nebrija University, Madrid, Spain.

[207]University of Pennsylvania, Philadelphia, PA, United States.

[208]School of Informatics, University of Skövde, Skövde, Sweden.

[209]Department of Developmental Psychology and Socialization (DPSS), University of Padova, Italy.

[210]Padova Neuroscience Center (PNC), University of Padova, Italy.

[211]Catholic University of Eichstätt-Ingolstadt, Germany.

[212]Institute of Psychology, University of Bern, Bern, Switzerland.

[213]Department of Psychology, New York University, New York, NY 10003, United States.

[214]Department of Engineering Science, Oxford University, Oxford, Oxfordshire, United Kingdom.

[215]Centre for Neurotechnology, School of Psychology and Neuroscience, University of Glasgow, Glasgow, United Kingdom.

[216]Department of Psychology, Rey Juan Carlos University, Madrid, Spain.

[217]Rotterdam School of Management, Erasmus University, The Netherlands.

[218]Department of Biological Psychology of Decision Making, Institute of Experimental Psychology, Heinrich Heine University Düsseldorf, Düsseldorf, Germany.

[219]Ming Hsieh Department of Electrical and Computer Engineering, University of Southern California, Los Angeles, CA, United States.

[220]New York University Abu Dhabi, United Arab Emirates.

[221]National Acoustic Laboratory, Sydney, Australia.

[222]Institute of Medical Psychology and Behavioural Neurobiology, Eberhard Karls University of Tübingen, Germany.

[223]CHUV-Centre Hospitalier Universitaire Vaudois, Service des Troubles du Spectre de l'Autisme et apparentés, Lausanne University Hospital, Les Allières – Av. Beaumont 23, 1011, Lausanne, Switzerland.

[224]Department of Developmental Psychology and Socialisation (DPSS), University of Padova, Italy.

[225]School of Psychology, University of Auckland, New Zealand.

[226]CHU Sainte-Justine, Canada.

[227]School of Psychological Sciences and Sagol School of Neuroscience, Tel Aviv University, Tel Aviv, Israel.

[228]School of biomedical engineering,University of British Columbia, Vancouver, Canada.

[229]Department of Psychology, University of Essex, United Kingdom.

[230]School of Biomedical Engineering, Sun Yat-sen University, China.

[231]Aging Research Center, Department of Neurobiology, Care Sciences, and Society, Karolinska Institutet and Stockholm University, Stockholm, Sweden.

[232]Psychological & Brain Sciences, Indiana University, Bloomington, IN, United States.

[233]Neuroimaging Group, Cajal Neuroscience Center, CSIC, Madrid, Spain.

[234]Department of Psychology and Program in Neuroscience, Bowdoin College, Brunswick, United States.

[235]Bruno Kessler Foundation, Italy.

[236]Department of Neuroscience, Georgetown University Medical Center, Washington, DC 20007, United States.

[237]Interdisciplinary Program in Bioengineering, College of Engineering, Seoul National University, Seoul, South Korea.

[238]Department of Liberal Arts and Social Sciences, University of Technology Nuremberg, Nuremberg, Germany.

[239]Dept. of Neuropsychology, Institute of Cognitive Neuroscience, Ruhr-Universität Bochum, Bochum, Germany.

[240]Institute of Behavioural Neuroscience, Department of Experimental Psychology, Division of Psychology and Language Sciences, University College London, United Kingdom.

[241]Differential Psychology and Psychological Assessment, University of Hamburg, Germany.

[242]Clinical Neurological Science Department, Schulich School of Medicine and Dentistry, Western University, London, Ontario, Canada.

[243]Department of Psychology, The University of Texas at Austin, Texas, United States.

[244]Department of Clinical and Biological Psychology, University of Bergen, Bergen, Norway.

[245]Danish Research Centre for Magnetic Resonance, Department of Radiology and Nuclear Medicine, Copenhagen University Hospital - Amager and Hvidovre, Copenhagen, Denmark.

[246]Faculty of Psychology, University of Vienna, Austria.

[247]Facultad de Ciencias de la Rehabilitación y Calidad de Vida, Universidad San Sebastián, Chile.

[248]Department of Psychology, University of Calgary, Calgary, Canada.

[249]Department of Linguistic, Literary and Aesthetic Studies, University of Bergen, Bergen, Norway.

[250]Center for Excellence in Brain Science and Intelligence Technology, Chinese Academy of Sciences, China.

[251]Wyss Center for Bio and Neuroengineering, Geneva, Switzerland.

[252]ABILITY Neurotech, Geneva, Switzerland.

[253]Philosophisch-pädagogische Fakültät, Katholische Universität Eichstätt-Ingolstadt, Eichstätt, Germany.

[254]Fakultät Gesellschaft und Ökonomie, Hochschule Rhein-Waal, Cleves, Germany.

[255]Institut für Psychologie, Universität Klagenfurt, Klagenfurt am Wörthersee, Austria.

[256]Ernst Strüngmann Institute (ESI) gGmbH of the Max Planck Society, Frankfurt, Germany.

[257]Faculty of Behavioural and Human Movement Sciences, Vrije Universiteit Amsterdam, The Netherlands.

[258]Otto-von-Guericke-University Magdeburg, Germany.

[259]Department of Human Development and Family Studies, The Pennsylvania State University, University Park, PA, United States.

[260]Department of Psychology V, Differential Psychology, Personality Psychology, and Psychological Diagnostics, Julius-Maximilians University of Würzburg, Germany.

[261]Department of Cognitive Science & Artificial Intelligence, Tilburg School of Humanities and Digital Sciences, Tilburg University, Tilburg, the Netherlands.

[262]Department of Systems Neuroscience, University Medical Center Hamburg-Eppendorf, Hamburg, Germany.

[263]Institut für Psychologie, Otto-von-Guericke-University Magdeburg, Magdeburg, Germany.

[264]Centre for Human Brain Health, University of Birmingham, Birmingham, United Kingdom.

[265]UiT Center for Language, Brain and Learning (C-LaBL), the Arctic University of Norway, Tromso, Norway.

[266]Linguistics and English Language, Lancaster University, Lancaster, United Kingdom.

[267]Nebrija Research Center in Cognition, Nebrija University, Madrid, Spain.

[268]Faculty of Psychology, UniDistance Suisse, Brig, Switzerland.

[269]School of Psychology and Humanities, University of Lancashire, United Kingdom.

[270]Sagol School of Neuroscience, Tel Aviv University, Tel Aviv, Israel.

[271]Department of Biosciences and Bioengineering, Indian Institute of Technology Bombay, India.

[272]Institute for Visualization and Interactive Systems (VIS), University of Stuttgart, Stuttgart, Germany.

[273]Engagement & Research Services, Erasmus University Rotterdam, The Netherlands.

[274]Institute for Globally Distributed Open Research and Education (IGDORE), Sweden.

[275]Universitätsklinikum Jena, Germany.

[276]School of Philosophy, Psychology and Language Sciences, University of Edinburgh, Edinburgh, United Kingdom.

[277]Department of Psychology, TUD Dresden University of Technology, Dresden, Germany.

[278]Concordia Institute for Information Systems Engineering (CIISE), Concordia University, Montreal, Canada.

[279]Department of Psychology, Carl von Ossietzky Universität Oldenburg, Germany.

[280]Linköping University, Sweden.

[281]Center for Mathematics, Computing and Cognition, Federal University of ABC (UFABC), São Paulo, Brazil.

[282]Center for Simulation Science, University of Stuttgart, Germany.

[283]University of Alabama, United States.

[284]Psychology Department, Princeton University, Princeton, NJ, United States.

[285]University College Dublin, Ireland.

[286]School of Psychological Sciences, Macquarie University, Sydney, Australia.

[287]Immanuel Kant Baltic Federal University, Russian Federation.

[288]Computational Neurology Group, Ruhr University Bochum, Bochum, Germany.

[289]University Catholic of Louvain (UCLouvain), Brussels, Belgium.

[290]Department of Psychology, Educational Science and Human Movement, University of Palermo, Italy.

[291]Institute of Psychology, Goethe University Frankfurt, Germany.

[292]Max-Planck-Institut für empirische Ästhetik, Germany.

[293]University of Manchester, Manchester, United Kingdom.

[294]Department of Neurology, University Clinic, Otto-von-Guericke University Magdeburg, Germany.

[295]Department of Biosciences and Bioengineering, Indian Institute of Technology Bombay, Mumbai, India.

[296]Department of Health Technology and Informatics, The Hong Kong Polytechnic University, Hong Kong.

[297]University of Lübeck, Germany.

[298]Goethe University Frankfurt, Germany.

[299]Programa de Investigación Asociativa (PIA) en Ciencias Cognitivas, Centro de Investigación en Ciencias Cognitivas (CICC), Facultad de Psicología, Universidad de Talca, Talca, Chile.

[300]University of Tübingen, Germany.

[301]Department of Neuroscience, University of Padova, Padova, Italy.

[302]Department of Neuroscience and Biomedical Engineering, Aalto University, Espoo, Finland.

[303]Department of Management, Hong Kong Baptist University, Hong Kong.

[304]Research Center for Psychological Science, Faculty of Psychology, University of Lisbon, Lisbon, Portugal.

[305]KESS Kindliche Entwicklung und Sprache Stärken, Offenbach, Germany.

[306]Department of Psychology, University of Girona, Girona, Spain.

[307]Department of Biomedical, Metabolic and Neural Sciences, University of Modena and Reggio Emilia, Italy.

[308]University of Hamburg, Germany.

[309]Department of Psychology and Behavioral Sciences, Zhejiang University, Hangzhou, China.

[310]Language and Genetics Department, Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands.

[311]Ernst Strüngmann Institute (ESI) for Neuroscience in Cooperation with the Max Planck Society, Frankfurt, Germany.

[312]Social Science Laboratory of Reading and Development in Children and Adolescents, Ministry of Education, & Center for Studies of Psychological Application, School of Psychology, South China Normal University, China.

[313]Centre for Mental Health and Brain Sciences, Swinburne University of Technology, Melbourne, Australia.

[314]University of Nevada, Reno, United States.

[315]Department of Health, Medical, and Neuropsychology, Leiden University, Leiden, The Netherlands.

[316]Leiden Institute for Brain and Cognition, Leiden University, The Netherlands.

[317]Faculty of Biology, Shenzhen MSU-BIT University, China.

[318]Graduate School of Health Sciences, Istanbul University, Istanbul, Turkey.

[319]Academy of Wellness and Human Development,Hong Kong Baptist University, Hong Kong.

[320]Human-AI experience team, JetBrains Research, Belgrade, Serbia.

[321]Bioeconomy division, RISE Research Institutes of Sweden, Stockholm, Sweden.

[322]Epilepsy Unit, Department of Clinical Neurosciences, University Hospitals and Faculty of Medicine, University of Geneva, Geneva, Switzerland.

[323]Department of Social and Behavioural Sciences, City University of Hong Kong, Hong Kong.

[324]Institute of Brain and Psychological Sciences, Sichuan Normal University, Chengdu, China.

[325]School of Psychological Sciences, Macquarie University, Australia.

[326]School of Sport Training, Chengdu Sport University, Chengdu, China.

[327]Center for Cognition and Brain Disorders/ Department of Neurology, The Affiliated Hospital of Hangzhou Normal University, Hangzhou, Zhejiang Province, China.

[328]Oxford Centre for Integrative Neuroimaging, FMRIB, Nuffield Department of Clinical Neurosciences, University of Oxford, United Kingdom.

[329]CAMH-University of Toronto, Canada.

[330]Centre for Chiropractic Research, New Zealand College of Chiropractic, Auckland, New Zealand.

[331]Donders Institute for Brain, Cognition and Behaviour, Radboud University Medical Center, Nijmegen, The Netherlands.

[332]Faculty of Information Technology, University of Jyväskylä, Jyväskylä, Finland.

[333]CLLE, Université Toulouse Jean Jaurès, CNRS, Toulouse, France.

[334]Department of Language Science, University of California, Irvine, United States.

[335]ELTE, Eotvos Lorand University, Budapest, Hungary.

[336]Radboud University Medical Center, Donders Centre for Medical Neuroscience, The Netherlands.

[337]Centre de recherche Cerveau et Cognition, Paul Sabatier University, Toulouse, France.

[338]Swartz Center for Computational Neuroscience, Institute for Neural Computation, University of California San Diego, La Jolla, California, United States.

[339]Department of Economics, Stockholm School of Economics, Stockholm, Sweden.

[340]Roche Pharma Research and Early Development, Neuroscience and Rare Diseases, Roche Innovation Center Basel, F. Hoffmann–La Roche Ltd., Basel, Switzerland.

[341]Department of Economics, University of Innsbruck, Innsbruck, Austria.

[342]Laboratory for Neurocognition and Applied Cognition, Faculty of Philosophy, University of Belgrade, Belgrade, Serbia.

[343]Department of Psychology, University of Alabama at Birmingham, United States.

[344]Neurobiology Research Unit, Copenhagen University Hospital, Rigshospitalet, Copenhagen, Denmark.

[345]Department of Psychology, Stanford University, Stanford, CA, United States.

[346]Department of Psychological & Brain Sciences, Indiana University, Bloomington, United States.

[347]School of Neurobiology, Biochemistry & Biophysics, Faculty of Life Sciences, Tel Aviv University, Tel Aviv, Israel.

[348]ESMT Berlin, Berlin, Germany.

[349]Institute of Operations and Decision Sciences, Corvinus Institute for Advanced Studies (CIAS), Corvinus University of Budapest, Budapest, Hungary.

[350]Institute of Psychology, ELTE Eötvös Loránd University, Budapest, Hungary.

[351]Department of Organizational Behavior, INSEAD, Singapore.

Funding Acknowledgements

The EEGManyPipelines project is supported by grants from German Research Foundation (DFG; BU 2400/11-1) to Niko A. Busch, by the DFG priority program "META-REP: A Meta-scientific Programme to Analyse and Optimise Replicability in the Behavioural, Social, and Cognitive Sciences" to Niko A. Busch and Elena Cesnaite, and Riksbankens Jubileumsfond (grant number: P21-0384) to Gustav Nilsonne, together with Mikkel C. Vinding and Niko A. Busch. Annalisa Pascarella acknowledges the Gruppo Nazionale Calcolo Scientifico-Istituto Nazionale di Alta Matematica (GNCS-INdAM). Darinka Trübutschek was supported by the European Union's Horizon 2020 research and innovation program under the Marie Sklodowska-Curie grant number 101023805. Alessandra Anzolin was funded by National Institutes of Health (NIH), National Center for Complementary and Integrative Health (NCCIH), Award Number R21-AT013513 and the Logan University-Martinos Center-Spaulding Integrative Neuroimaging Training Fellowship. Yana Arkhipova was supported by the German Research Foundation DFG via SFB 1287, project B03. Maria Azanova was supported by the German Federal Ministry of Education and Research (BMBF) and the Max Planck Society. Ceren Battal was supported by The Portuguese Foundation for Science and Technology (FCT). Giulio Bernardi was supported by an ERC Starting Grant (TweakDreams; Grant agreement ID: 948891). Claire Braboszcz was funded by the European Union's Horizon 2020 research and innovation programme under Marie Skłodowska-Curie grant agreement No. 801342 (Tecniospring INDUSTRY) and the Government of Catalonia's Agency for Business Competitiveness (ACCIÓ). Florian Bublatzky was funded by Deutsche Forschungsgemeinschaft (DFG, German Research Foundations) – Project-ID 290032019 and 566299199. Xim Cerda-Company was supported by a project (PID2023-151083NA-I00) funded by MICIU/AEI/10.13039/501100011033 and FEDER, UE. Karan Chugani was funded by the European Union's Horizon 2020 research and innovation programme under grant agreement No. 945307 (Emotional Cities). Oren Civier, Matthew Hughes, Tom Johnstone, William Woods and David White acknowledge the facilities and scientific and technical assistance of the National Imaging Facility, a National Collaborative Research Infrastructure Strategy (NCRIS) capability, at Swinburne Neuroimaging, Swinburne University of Technology. César E. Corona-González expresses sincere gratitude to SECIHTI for the financial support received during the PhD program (reference number: 894658). Martin Dietz was supported by the Lundbeck Foundation (ref. R322-2019-2711). M. Dolores del Castillo was funded by the MICIU/AEI/ 10.13039/501100011033 (Spanish Ministry of Science, Innovation and Universities, Spanish State Research Agency; grant PID2023-150271NB-C21), and iRoboCity2030-CM (grant TEC-2024/TEC-62). Anna Dreber was funded by the Knut and Alice Wallenberg Foundation. Linda Drijvers was funded by the Max Planck Society through a Minerva Fast Track Fellowship awarded to Linda Drijvers. Brandi Lee Drisdelle was funded by the Natural Sciences and Engineering Research Council of Canada (NSERC) Postdoctoral Fellowships program and the Fonds de recherche du Québec – Nature et technologies (FRQNT) Postdoctoral research scholarship program. Gian Marco Duma was funded by Ricerca Corrente from Italian Ministry of Health. Guillaume Dumas was funded by the Institute for Data Valorization, Montreal (IVADO; CF00137433), the Fonds de recherche du Québec (FRQ; 285289), the Natural Sciences and Engineering Research Council of Canada (NSERC; DGECR-2023-00089), the Canadian Institute for Health Research (CIHR 192031; SCALE). The computational analyses were enabled in part by support provided by Calcul Québec

(www.calculquebec.ca) and Digital Research Alliance of Canada (www.alliancecan.ca). Benedikt Ehinger was funded by Deutsche Forschungsgemeinschaft (DFG, German Research Foundations) in the Emmy Noether Programme - Project-ID 538578433. Stefanie Evas was funded by an Industry PhD award comprising an Australian Government Research Training Program Stipend (RTPS) administered by the University of Adelaide, together with funding from the Commonwealth Scientific and Industrial Research Organisation (CSIRO, Australia) and Fonterra Co-operative Group Limited (New Zealand). Gordon B. Feld was funded by Deutsche Forschungsgemeinschaft (DFG, German Research Foundations) in the Emmy Noether Programme - 404679789. Norman Forschack was supported by the Deutsche Forschungsgemeinschaft (DFG grants: FO 1615/1-1, MU 972/29-1). Haydée G. García Lázaro was supported by the Smith-Kettlewell Eye Research Institute's C.V. Starr Fellowship Fund and the E. Matilda Ziegler Foundation for the Blind. Carlos González-García was supported by Project PID2023-149428NB-I00 funded by MCIN/AEI/10.13039/501100011033 and by FEDER, EU, and Grant RYC2021-033536-I funded by MCIN/AEI/10.13039/501100011033 and by the European Union NextGeneration EU/PRTR. Benjamin J. Griffiths was funded by a Leverhulme Trust Early Career Fellowship (ECF-2021-628). Anthony M. Harris was supported by the Australian Research Council (DE220101019). Stefan Haufe was supported by an ERC Starting Grant (TrueBrainConnect; Grant agreement ID: 758985). Alice Hodapp was funded by the German Research Foundation DFG Emmy Noether grant RA 2715/2-1 awarded to Milena Rabovsky. Chun-Hsien Hsu was funded by National Science and Technology Council, Taiwan (MOST-108-2636-H-008-001 & 114-2410-H-008 -042). Christoph Huber-Huber was supported by a Marie Skłodowska-Curie Individual Fellowship (No. 846392) and by the Ministry of Universities and Research Italy (MUR) within the NextGenerationEU/PNRR framework (grant number 0000027). Agustin Ibanez is supported by grants from the Multi-partner consortium to expand dementia research in Latin America [ReDLat, supported by Fogarty International Center (FIC), National Institutes of Health, National Institutes of Aging (R01 AG057234, R01 AG075775, R01 AG21051, R01 AG083799, CARDS-NIH, R01 AG057234), Alzheimer's Association (SG-20-725707), Rainwater Charitable Foundation – The Bluefield project to cure FTD, and Global Brain Health Institute)], ANID/FONDECYT Regular (1250091 and 1210176 and 1220995); ANID/PIA/ANILLOS ACT210096; JPI JPND-Care, DISCeRN 2025 - Health and Social Care Research with a Focus on the Moderate and Late Stages of Neurodegenerative Diseases; FONDEF ID20I10152, and ANID/FONDAP 15150012; Wellcome Trust award for BRAIN-CLIMA: Investigating the Combined Impact of Heat and Air Pollution on Blood-Brain Barrier Integrity and Brain Aging in Latin America, (335293/Z/25/Z), and the CliCBrain (Horizon ID: 101236426; DOI 10.3030/101236426, Marie Skłodowska-Curie Actions - MSCA) . The contents of this publication are solely the responsibility of the authors and do not represent the official views of these institutions. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript. Bradley N. Jack was funded by the Australian Research Council (DE220100739). Daniel Kaiser was supported by an ERC starting grant (PEP, ERC-2022-STG 101076057) and by the DFG under Germany's Excellence Strategy (EXC 3066/1 "The Adaptive Mind", project number 533717223). Taeho Kang was partially funded by the Institute of Information & Communications Technology Planning & Evaluation (IITP) Grants from the Korean government (No. 2017-0-00451), partially funded by the Austrian Science Fund (FWF) 10.55776/I6222. Hamid Karimi-Rouzbahani was supported by Newton International Fellowship (NIF\R1\192608) and follow-on

fundings (AL231037 & AL24100035) from the Royal Society and Australian Research Council DECRA Fellowship DE230100608. Bruno Kopp was funded by Karlheinz-Hartmann Stiftung, Hannover. Daniel S. Kluger was funded by the European Union (ERC StG No. 101162169) and the Innovative Medical Research (IMF) programme of the University of Münster (grant ID KL 1 2 22 01). Corinna Kührt was funded by the German Research Foundation (DFG, Deutsche Forschungsgemeinschaft) as part of the Collaborative Research Centre “Volition and Cognitive Control” – SFB 940/2 – and Germany’s Excellence Strategy – EXC 2050/1 – Project ID 390696704 – Cluster of Excellence “Centre for Tactile Internet with Human-in-the-Loop” (CeTI) of Technische Universität Dresden. Martin Lamoš was supported by the core facility MAFIL (MEYS CR, LM2023050 Czech-BioImaging) and by the European Union’s Horizon Europe research and innovation programme under grant agreement No 101099481 (RELIEVE). Vladimir Litvak was supported by core funding from Wellcome Centre for Human Neuroimaging (203147/Z/16/Z). Eduardo López-Caneda was supported by the Grant RYC2024-049840-I, funded by MICIU/AEI/10.13039/501100011033. Alicia Luque was supported by the César Nombela Talent Attraction Grant (2023-T1/PH-HUM-29098), funded by the General Directorate of Research and Technological Innovation of the Regional Government of Madrid (Spain). Antonio Maffei is funded by an individual SID grant from the Department of Devolpmental Psychology and Socialization (MAFF_BIRD24_01). Radek Marecek was supported by the core facility MAFIL supported by MEYS CR (LM2023050 Czech-BioImaging), part of the Euro-BioImaging (www.eurobioimaging.eu) ALM and Medical Imaging Node (Brno, CZ). Anna Marzecova was funded by the Deutsche Forschungsgemeinschaft (DFG) grant JO-787/6-1 and Research Foundation Flanders (FWO) postdoctoral fellowship (12V5620 N). Jason B. Mattingley was funded by an Investigator Grant awarded by the Australian National Health and Medical Research Council (NHMRC, GNT2010141). Sara Mazzini was funded by the Max Planck Society through a Minerva Fast Track Fellowship awarded to Linda Drijvers. Takfarinas Medani, Raymundo Cassani and Juan Garcia-Prieto were supported by the National Institute of Biomedical Imaging and Bioengineering (NIBIB) of the National Institutes of Health (NIH) under award numbers R01EB026299. Maria Montefinese was funded by the Investment line 1.2 “Funding projects presented by young researchers” (CHILDCONTROL) from the European Union - NextGenerationEU. Faisal Mushtaq was supported by funding from UKRI BBSRC (BB/X008428/1), the National Institute for Health and Care Research (NIHR) Leeds Biomedical Research Centre (NIHR203331) and the EEG101 European COST Action (CA24148). Guiomar Niso was supported by NIH NIBIB R01EB030896, INDoS European COST Action (CA24161) and by the Spanish Government RYC2021-033763-I and PID2023-150034OA-I00. Norberto E. Naal-Ruiz was supported by Tecnologico de Monterrey and the Secretariat of Science, Humanities, Technology, and Innovation (SECIHTI). Robert Oostenveld was funded by the European Union through the SIESTA project (Horizon Europe) under Grant number 101131957. Pavel Prado was supported by ANID/FONDECYT Regular 1260530. Andrea B. Protzner was supported by a Natural Sciences and Engineering Research Council of Canada Discovery Grant (05299–2020). Milena Rabovsky was supported by the German Research Foundation DFG Emmy Noether grant RA 2715/2-1. Manuel Rausch was supported by the Deutsche Forschungsgemeinschaft (grant RA 2988/4-1). Pavel Riha was supported by the Czech Health Research Council AZV NW25-04-00345 and the core facility MAFIL supported by MEYS CR (LM2023050 Czech-BioImaging), part of the Euro-BioImaging (www.eurobioimaging.eu) ALM and Medical Imaging Node (Brno, CZ).

Johannes Rodrigues was funded by Deutsche Forschungsgemeinschaft (DFG, German Research Foundations) – Project-ID 540631605 and the European Union through the project “Hyperautomation Ökosystem (Hyko)“ (1.2-StMWKF.4-UFR-002) in the European fonds for regional development (ERDF). Judith Schepers was supported by the International Max Planck Research School for Intelligent Systems (IMPRS-IS). Noor Seijdel was funded by the Max Planck Society through a Minerva Fast Track Fellowship awarded to Linda Drijvers. Magdalena Senderecka was supported by the National Science Centre of Poland (2020/38/E/HS6/00490). J. Ignacio Serrano was supported by the MICIU/AEI/ 10.13039/501100011033 (Spanish Ministry of Science, Innovation and Universities, Spanish State Research Agency; grant PID2023-150271NB-C21), and iRoboCity2030-CM (grant TEC-2024/TEC-62). Cassie Ann Short was supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundations) grant WA2593/9-1. Esaú V. Sirius was funded by São Paulo Research Foundation (FAPESP) (grant number 2019/06423-7). René Skukies would like to thank the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) for supporting this work by funding - EXC2075 – 390740016 under Germany's Excellence Strategy. We acknowledge the support by the Stuttgart Center for Simulation Science (SimTech). Aureli Soria-Frisch was funded by the European Union’s Horizon 2020 research and innovation programme under grant agreement No. 945307 (Emotional Cities). Simon R. Steinkamp was funded by a Novo Nordisk Foundation Exploratory Interdisciplinary Synergy Grant, ref NNF20OC0064869. Anđela Šoškić and Vanja Ković were supported by Ministry of Science, Innovation, and Technological Development (contract no. 451-03-33/2026-03/ 200163). Alessandro Tavano was supported by the Deutsche Forschungsgemeinschaft (DFG) Grant number 510229904 and the Max Planck Society. Nivethida Thirugnanasambandam was supported by DBT/Wellcome Trust India Alliance Fellowship (IA/CPHI/16/1/50264)). Luca Turella was supported by the European Union – Next Generation EU, in the framework of the project MNESYS - A MULTISCALE INTEGRATED APPROACH TO THE STUDY OF THE NERVOUS SYSTEM IN HEALTH AND DISEASE (Prot. n. PE00000006). Luca Turella was also supported by the European Union - Next Generation EU PRIN 2022 (DD 104 02/02/22) – PNRR – M4 – C2 – INV1.1 – PRIN –Project Title [COMPArative charACTterization of the parietal lobe: a multimodal approach (COMPAACT project)] – project identifier code [2022JT852C] – CUP [E53D23022180006]. The views and opinions expressed are solely those of the authors and do not necessarily reflect those of the European Union, nor can the European Union be held responsible for them. Eric Luis Uhlmann was supported by funding from the INSEAD R&D Committee. José Luis Ulloa was supported by the Programa de Investigación Asociativa (PIA) en Ciencias Cognitivas of the Cognitive Science Research Center at Universidad de Talca and the FONDEQUIP-ANID project EQM190153. Adrià Vilà-Balló was funded by MICIU/AEI (10.13039/501100011033), FEDER, EU (project PID2023-151083NA-I00), the Dirección General de Ordenación del Juego, Ministerio de Derechos Sociales, Consumo y Agenda 2030 (project SUBV24/00014), and the European Union’s Horizon Europe programme (grant agreement No. 101155881; HORIZON-HLTH-2024-DISEASE-03-two-stage). Marijn van Vliet was funded by the Research Council of Finland, (RCF) grant #343385. Ingmar E. J. de Vries was supported by a Marie Skłodowska-Curie Postdoctoral Fellowship (No. 101060807). Gwen van der Wijk was supported in her work through a Mamdani Family Foundation Graduate Scholarship and Alberta Graduate Excellence Scholarship (AGES: International). Audrey Wong-Kee-You was supported by the Rachel C. Atkinson Postdoctoral Fellowship. Judy D. Zhu was

supported by the Australian National Imaging Facility, a National Collaborative Research Infrastructure Strategy (NCRIS) capability, at the KIT-Macquarie Brain Research (MEG) Laboratory, Macquarie University.

Competing interests

Omid Abbasi is the CEO of Virgobit GmbH, a company that creates healthcare software solutions. Paulo Rodrigo Bazán started working at Brain Products GmbH in March 2024. Martina Berto started working at Neurotechnology UAB (BrainAccess) in October 2024. Karan Chugani, Aureli Soria-Frisch, and Claire Braboszcz are employees of Starlab, Barcelona. Maximilian Hommelsen is a full-time employee of F. Hoffmann-La Roche Ltd. Aureli Soria-Frisch works at Starlab Barcelona SL and is an advisor of Neuroelectrics. Mahmoud Hassan is the Founder and CEO of MINDIG. Jörg Hipp is a full-time employee of F. Hoffmann-La Roche Ltd. Agustin Ibanez is supported by the Multi-Partner Consortium to Expand Dementia Research in Latin America (ReDLat), supported by the Fogarty International Center (FIC), the National Institutes of Health, the National Institute on Aging (R01 AG057234, R01 AG075775, R01 AG21051, R01 AG083799, CARDS-NIH 75N95022C00031), the Alzheimer's Association (SG-20-725707), the Rainwater Charitable Foundation – The Bluefield Project to Cure FTD, and the Global Brain Health Institute. Agustin Ibanez is also supported by ANID/FONDECYT Regular (1250091, 1210176, 1220995) and ANID/FONDAP/15150012. Muhammad Samran Navid is associated with Brain Products GmbH since December 2025. All other authors declare no competing interests.

Author contributions

Current EEGManyPipelines Steering Committee constitution: Johannes Algermissen, Niko A. Busch, Elena Cesnaite, Nastassja L. Fischer, Claudia Gianelli, Gustav Nilsonne, Annalisa Pascarella, Darinka Trübutschek, Mikkel C. Vinding, Andrea Vitale, Yu-Fang Yang

Steering Committee members who initiated the project: Johannes Algermissen, Niko A. Busch, Nastassja L. Fischer, Claudia Gianelli, Josh Koen, Charlotte R. Marshall, Muhammad Samran Navid, Gustav Nilsonne, Annalisa Pascarella, Tuomas Puoliväli, Mehdi Senoussi, Darinka Trübutschek, Mikkel C. Vinding, Andrea Vitale, Yu-Fang Yang, Jeremy Yeaton

Drafting the manuscript: Johannes Algermissen, Niko A. Busch, Elena Cesnaite, Nastassja L. Fischer, Claudia Gianelli, Gustav Nilsonne, Annalisa Pascarella, Darinka Trübutschek, Mikkel C. Vinding, Andrea Vitale, Yu-Fang Yang, Eric L. Uhlmann

EEG dataset sent to the analysts (experiment design, data collection, and quality assurance): Niko A. Busch

Experimental design of the EEGManyPipelines project: Johannes Algermissen, Niko A. Busch, Elena Cesnaite, Nastassja L. Fischer, Claudia Gianelli, Charlotte R. Marshall, Muhammad Samran Navid, Gustav Nilsonne, Annalisa Pascarella, Tuomas Puoliväli, Darinka Trübutschek, Mikkel C. Vinding, Andrea Vitale, Yu-Fang Yang

Analysis team (recruitment, point of contact, and data management): Johannes Algermissen, Niko A. Busch, Elena Cesnaite, Nastassja L. Fischer, Claudia Gianelli, Charlotte R. Marshall, Gustav Nilsonne, Annalisa Pascarella, Tuomas Puoliväli, Darinka Trübutschek, Mikkel C. Vinding, Andrea Vitale, Yu-Fang Yang

Analysis team (data analysis and visualisation): Johannes Algermissen, Elena Cesnaite, Nastassja L. Fischer, Annalisa Pascarella, Mikkel C. Vinding, Andrea Vitale

Participated as members of the advisory board and reviewed and edited the manuscript: Balazs Aczel, Mike X. Cohen, Arnaud Delorme, Anna Dreber, Jörg Hipp, Felix Holzmeister, Magnus Johannesson, Vanja Ković, Robert Oostenveld, Yuri G. Pavlov, Cyril Pernet, Russell Poldrack, Aina Puce, Tom Schonberg, Martin Schweinsberg, Anđela Šoškić, Barnabas Szaszi, Eric L. Uhlmann

Participated as members of analysis teams and reviewed and edited the manuscript: Omid Abbasi, Timofey Adamovich, Lena Adel, Nikita Agarwal, Blanca Aguado-López, Uma Ajmeria, Esra Al, Hannah Aldeen, Ricardo J. Alejandro, Evgeniia A. Alenina, Meera Alex, Ettore Ambrosini, Douglas J. Angus, Alessandra Anzolin, Stefan Appelhoff, Giorgio Arcara, Yana Arkhipova, Maria Azanova, Chiara Bagattini, Daniel H. Baker, Christoph Bamberg, Francesca M. Barbero, Ceren Battal, Anna-Katharina R. Bauer, Burcu Bayram, Paulo R. Bazán, Henry Beale, Elena Belanova, Zara M. Bergström, Giulio Bernardi, Martina Berto, Monica Betta, Isabelle Blanchette, Marta Bortoletto, Davide Bottari, Claire Braboszcz, Daniel E. Bradford, Joseph L. Brooks, Chloe M. Brunskill, Florian Bublatzky, Daniel Büchel, Roberta P. Calce, Mariagrazia Capizzi, Raymundo Cassani, María Concepción Castellanos, Luigi Cattaneo, Xim Cerda-Company, Lixiang Chen, Xinyuan Chen, Yiyu Chen, Yoojeong Choo, Hu Chuan-Peng, Karan Chugani, Radoslaw Martin Cichy, Barbora Cimrová, Simon Ciranka, Oren Civier, Carlotta Cogoni, Michel-Pierre Coll, Julie Coloigner, Angela Conejero, Martin Constant, César E. Corona-González, Antonio Criscuolo, Damian Cruse, Claire Cury, Artur Czeszumski, Martin J. Dahl, Matthew Davidson, Ingmar E. J. de Vries, M. Dolores del Castillo, Vincent DeLuca, Gianpaolo Demarchi, Kobe Desender, Marcelo Dias, Darcy Waller, Martin Dietz, Olaf Dimigen, Hao Ding, Li Dong, Linda Drijvers, Brandi Lee Drisdelle, Gian Marco Duma, Guillaume Dumas, Benedikt Ehinger, Alexandra K. Emmendorfer, Alexander Enge, Stefanie Evas, Tommaso Fedele, Alessandra Federici, Gordon B. Feld, Lauren K. Fink, Shai Fischer, Emilia Fló, Silvia Formica, Norman Forschack, Bettina Forster, Marcel Franz, Léon Franzen, Cassidy M. Fry, Alejandro Galvez-Pol, Haydee G. García-Lázaro, José Carlos García Alanis, Patricia Garrido-Vásquez, Billy Gerdfeldter, Georgia Gerike, Magdalena Gippert, Carlos González-García, Dario Gordillo, Anna Grabowska, Lisa-Marie Greenwood, Benjamin J. Griffiths, Demetrio Grollero, Joachim Gross, Danièle A. Gubler, Pierre Guilleminot, Berna Güler, Christopher Gundlach, Eren Günseli, Mingqian Guo, Malte R. Güth, Céline Haciahmet, Itay Hadas, Jarmo Hämäläinen, Sizhu Han, Anthony M. Harris, Mahmoud Hassan, Stefan Haufe, Anibal S. Heinsfeld, Peer Herholz, Mario Hervault, Aron T. Hill, Alice Hodapp, Maximilian Hommelsen, Hanna Honcamp, Thomas J. Hosang, Chun-Hsien Hsu, Tzu-Yu Hsu, Changrun Huang, Christoph Huber-Huber, Matthew Hughes, Agustin Ibanez, Behzad Iravani, Kylie Isenburg, Bradley N. Jack, Mina Jamshidi Idaji, Oskar H. Jefsen, El Jeong, Fang Jiang, Ann-Kathrin Joechner, Matthew R. Johnson, Ki-Young Jung, Neda Kaboodvand, Daniel Kaiser, Patrycja Kałamała, Taeho Kang, Nikolai Kapralov, Hamid Karimi-Rouzbahani, Mikolaj Kegler,

Nicholas J. Kelley, Simon Kern, Casper Kerrén, Songhee Kim, Alina Kiseleva, Laura-Isabelle Klatt, Felix Klotzsche, Daniel S. Kluger, Peter König, Bruno Kopp, Julian Q. Kosciessa, Maciej Kosilo, Boris Kotchoubey, Layla Kouara, Sander Krewinkel, Leon O. H. Kroczek, Corinna Kührt, Sara Lago, Martin Lamoš, Nicolas Langer, Dmitry Lazurenko, Pierre Le Denmat, Ty Lees, Tim Lehmann, Tomas Lenc, Cédric Lenoir, Wanru Li, Zefeng Li, Anna M. Liesefeld, Heinrich R. Liesefeld, Maxim Likhanov, Evan N. Lintz, Giulia Lioi, Vladimir Litvak, Xiaoyi Liu, Yikang Liu, Yu-Hui Lo, Eduardo López-Caneda, David López-García, Corinna Lorenz, Andreas Löw, Runhao Lu, Zitong Lu, Alicia Luque, David Lydon-Staley, Oskar MacGregor, Antonio Maffei, Martin E. Maier, Lisa M. Makowski, Francesco Mantegna, Eleonora Marcantoni, María Carmen Martín-Buro, Ana Martinovici, Anna Marzecova, Fabio Masina, Jason B. Mattingley, Sara Mazzini, Takfarinas Medani, Ting Mei, David Melcher, Qingqing Meng, Nick S. Menger, Ahmad Mheich, Milan Mitka, Claire Monroy, Maria Montefinese, David Moreau, Quentin Moreau, Liad Mudrik, Faisal Mushtaq, Nicholas E. Myers, Norberto E. Naal-Ruiz, Foroogh Najafi, Hong-Viet V. Ngo, Lu Nie, Martin Nilsson, Guiomar Niso, Erika Nyhus, Joyce Oerlemans, Emanuele Olivetti, Joan Orpella, Eduard Ort, Ana F. Palenciano, Kyoungeun Park, Bernhard Pastötter, Zita Patai, Katharina Paul, Artur José M. Paulo, Giovanni Pellegrino, Sergio M. Pereira Soares, Franco Pestilli, Rachele Pezzetta, Daniela M. Pfabigan, Anna Plachti, Ulrich Pomper, Tzvetan Popov, Pavel Prado, Andrea B. Protzner, Yanina Prystauka, Haoyue Qian, Milena Rabovsky, Janir Ramos da Cruz, Manuel Rausch, Amit Rawal, Stefan Repplinger, Danielle R. Rice, Igor Riecansky, Pavel Riha, Johannes Rodrigues, M. Paula Roncaglia, Michael Rose, Marlene Rösner, Jason Rothman, Simon Ruch, Philipp Ruhnau, Adi Sarig Golan, Arkaprovo Sarkar, Akul Satish, Cristina Scarpazza, Christoph Scheffel, Judith Schepers, Antonio Schettino, Barbara Schmidt, Fabian Schmidt, Anna-Lena Schubert, Anna-Lisa Schuler, Catriona L. Scrivener, Magdalena Senderecka, Ulrike Senftleben, Fatih Serin, J. Ignacio Serrano, Mehran Shabanpour, Veronika Shamova, Idris Shareef, Dmitry Sherbina, Cassie Ann Short, Carine Signoret, Esaú V. P. Sirius, René Skukies, Aureli Soria-Frisch, Firat Soylu, Sebastian Speer, Tim Paul Steinfath, Robert Steinhauser, Simon R. Steinkamp, Tilman Stephani, Alexander Strobel, Dawid Strzelczyk, Pheobe Sun, Yanan Sun, Caroline Surrey, Nikolay Syrov, Marco Tagliaferri, Siddharth Talwar, Vincenza Tarantino, Alessandro Tavano, Jason R. Taylor, Carsten Thiele, Nivethida Thirugnanasambandam, Lara Todorova, Aleksandra Tomić, Stefan J. Troche, Lin-Yuan Tseng, Sarah Tune, Gözem Turan, Luca Turella, José Luis Ulloa, Michael Valiadis, Antonino Vallesi, Gwen van der Wijk, Marijn van Vliet, Marijn van Wingerden, Enrico Varano, Mohith M. Varma, Margarida Vasconcelos, Yoana Vergilova, Adrià Vilà-Balló, Antonino Visalli, Giada Viviani, Robin Vloeberghs, Jan Wacker, Ya-Jie Wang, Meng-Yun Wang, Tianlu Wang, Duan Wei, Markus Werkle-Bergner, David White, Simon Whitton, Kristina Wiebels, Stefan Wiens, Christoph A. Wittkamp, Maren-Isabel Wolf, Audrey Wong-Kee-You, Will Woods, Franz Wurm, Syanah C. Wynn, Jiushu Xie, Alba Xifra-Porxas, Weiyong Xu, Lev Yakovlev, Jinbiao Yang, Mustafa Yavuz, Vahab Youssofzadeh, Rongjun Yu, Talha Zafar, Sara Zago, Ilya Zakharov, Marta Z. Zakrzewska, Juanli Zhang, Zhen Zhang, Shanshan Zhen, Xinqi Zhou, Judy D. Zhu, Siyu Zhu, Qian Zhuang, Catharina Zich, Reza Zomorrodi, Camila Zugarramurdi.