# The Methodology and Dataset of the CoScience EEG-Personality Project - A Large-Scale, Multi-Laboratory Project Grounded in Cooperative Forking Paths Analysis

Paul, Katharina; Short, Cassie Ann; Beauducel, André; Carsten, Hannes Per; Härpfer, Kai; Hennig, Jürgen; Hewig, Johannes; Hildebrandt, Andrea; Kührt, Corinna; Mueller, Erik Malte; Munk, Aisha; Osinsky, Roman; Porth, Elisa; Riesel, Anja; Rodrigues, Johannes; Scheffel, Christoph; Stahl, Jutta; Strobel, Alexander; Wacker, Jan

Veröffentlichungsversion / Published Version
Zeitschriftenartikel / journal article

**Empfohlene Zitierung / Suggested Citation:**
Paul, K., Short, C. A., Beauducel, A., Carsten, H. P., Härpfer, K., Hennig, J., ... Wacker, J. (2022). The Methodology and Dataset of the CoScience EEG-Personality Project - A Large-Scale, Multi-Laboratory Project Grounded in Cooperative Forking Paths Analysis. *Personality Science*, *3*, 1-26. https://doi.org/10.5964/ps.7177

PERSONALITY Science

PsychOpen GOLD

Projects & Data

# The Methodology and Dataset of the CoScience EEG-Personality Project – A Large-Scale, Multi-Laboratory Project Grounded in Cooperative Forking Paths Analysis

Katharina Paul[1], Cassie Ann Short[1], André Beauducel[2], Hannes Per Carsten[1], Kai Härpfer[1], Jürgen Hennig[3], Johannes Hewig[4], Andrea Hildebrandt[5], Corinna Kührt[6], Erik Malte Mueller[7], Aisha Munk[3], Roman Osinsky[8], Elisa Porth[9], Anja Riesel[1], Johannes Rodrigues[4], Christoph Scheffel[6], Jutta Stahl[9], Alexander Strobel[6], Jan Wacker[1]

**[1]** *Institut für Psychologie, Universität Hamburg, Hamburg, Germany.* **[2]** *Institut für Psychologie, Rheinische-Friedrich-Wilhelms-Universität Bonn, Bonn, Germany.* **[3]** *Fachgebiet Psychologie, Justus-Liebig-Universität Gießen, Gießen, Germany.* **[4]** *Institut für Psychologie I, Julius-Maximilians-Universität Würzburg, Würzburg, Germany.* **[5]** *Department für Psychologie, Carl von Ossietzky Universität Oldenburg, Oldenburg, Germany.* **[6]** *Fakultät Psychologie, Technische Universität Dresden, Dresden, Germany.* **[7]** *Fachbereich Psychologie, Philipps-Universität Marburg, Marburg, Germany.* **[8]** *Institut für Psychologie, Universität Osnabrück, Osnabrück, Germany.* **[9]** *Fachbereich Psychologie, Universität zu Köln, Köln, Germany.*

Personality Science, 2022, Vol. 3, Article e7177, https://doi.org/10.5964/ps.7177

**Received:** 2021-07-20 • **Accepted:** 2021-12-23 • **Published (VoR):** 2022-06-08

**Handling Editor:** John F. Rauthmann, Universität Bielefeld, Bielefeld, Germany

**Reviewing:** *Round 1* - Mathias Benedek; Anonymous #1. Open reviews are available. [see Index of Supplementary Materials]

**Corresponding Author:** Cassie Ann Short, Universität Hamburg, Fakultät für Psychologie und Bewegungswissenschaft, Differentielle Psychologie und Psychologische Diagnostik, Von-Melle-Park 5, 20146 Hamburg, Germany. E-mail: Cassie.ann.short@uni-hamburg.de

**Supplementary Materials:** Materials, Preregistration [see Index of Supplementary Materials]

## Abstract

Despite a plethora of research, associations between individual differences in personality and electroencephalogram (EEG) parameters remain poorly understood due to concerns of low replicability and insufficiently powered data analyses due to relatively small effect sizes. The present article describes how a multi-laboratory team of EEG-personality researchers aims to

alleviate this unsatisfactory status quo. In particular, the present article outlines the design and methodology of the project, provides a detailed overview of the resulting large-scale dataset that is available for use by future collaborators, and forms the basis for consistency and depth to the methodology of all resulting empirical articles. Through this article, we aim to inform researchers in the field of Personality Neuroscience of the freely available dataset. Furthermore, we assume that researchers will generally benefit from this detailed example of the implementation of cooperative forking paths analysis.

## Keywords

personality neuroscience, EEG, replicability, multiverse analysis, cooperative forking paths

### Relevance Statement

This initial application of a cooperative Forking Path Analysis aims to increase replicability in Personality Neuroscience by including collaborator-reviewed hypotheses, sufficient statistical power, and a multiverse analysis of analysis choices.

### Key Insights

- Personality Neuroscience suffers from small sample sizes and analysis flexibility.
- Initial application of a cooperative forking path analysis to increase replicability.
- Detailed description of an existing dataset to encourage distribution.

It has become clear that low statistical power and undisclosed flexibility in data analysis are major drivers of the current replicability crisis (Open Science Collaboration, 2015; Simmons et al., 2011), to which psychology and cognitive neuroscience are not immune (Button et al., 2013; Open Science Collaboration, 2015; Szucs & Ioannidis, 2017). Both factors are exacerbated in electroencephalography (EEG) research (Luck & Gaspelin, 2017) and, more generally, in research aiming to link individual differences in neural responses to variations in personality. At the core of this issue are four problems: a fragmented theoretical landscape, modest sample sizes, undisclosed flexibility in data analysis, and private datasets and analysis scripts, which the CoScience EEG-personality project aims to address through the initial application of a novel approach for empirical research termed "cooperative Forking Path Analysis" (cFPA), described in detail by Wacker (2017).

First, *a fragmented theoretical landscape* exists within the literature, which may be more difficult to overcome given the limitations and difficulties of single-laboratory-based research. This has led to an emphasis in the research community on collaborative approaches, data sharing, and integration of findings across studies (e.g., through meta analyses). Within the cFPA approach, the combined expertise of teams of researchers helps to arrive at more refined and informed theoretical and practical conclusions through a collaborator-review of all research proposals and the overall research design.

Second, the abundance of studies with *modest sample sizes* impedes real scientific progress by reducing the statistical power to detect a true effect and increasing the likelihood that the effects reported are false positives. Studying individual differences by definition requires between-subjects designs. A recent report estimated a modest median correlation of $r$ = .19 for individual differences research (Gignac & Szodorai, 2016), and there is little reason to expect that associations between self-report measures and electrophysiological variables surpass this benchmark, especially since this estimate is likely inflated by publication bias and methodological overlap. Furthermore, a large sample size increases the ability to produce more stable and accurate correlations, especially for modest associations (Schönbrodt & Perugini, 2013). The multi-laboratory collaboration aspect of the cFPA approach increases the final sample size and, thus, statistical power, while avoiding added costs and time resources required of each laboratory.

Third, similar to functional magnetic resonance imaging (fMRI; Botvinik-Nezer et al., 2020), EEG data requires complex pre-processing and analysis routines with a number of equally defensible alternatives for every processing step resulting in a vast number of researcher degrees of freedom (Simmons et al., 2011). *Undisclosed flexibility in data analysis* is a possible source of false positive findings and could thereby hamper replicability. While there have been attempts to standardize EEG analysis routines (for example, Gabard-Durnam et al., 2018; Rodrigues et al., 2021), their application remains scarce because unique technical equipment, experimental set-ups or participant samples require more customized routines. Forking Path Analysis (Wacker, 2017) has been proposed as a new analytical tool in order to estimate the true impact of this flexibility, to contribute to the development of a standard, and to test the robustness of reported findings. In this framework, instead of basing a conclusion on one single and potentially arbitrarily derived pre-processing and analysis path, the outcome is evaluated against several defensible pre-processing and analysis decisions. However, the existence of several defensible pre-processing and analysis options raises the question on which of the defensible options conclusions should be based. In contrast to similar approaches like multiverse analysis (Steegen et al., 2016) or specification curve analysis (Simonsohn, Simmons, & Nelson, 2020), cFPA addresses this issue by means of collective identification of the most appropriate pre-processing and analysis path in addition to a great number of defensible alternative paths. Specifically, for each single step of the pre-processing and analysis, the team of researchers discusses and identifies as many defensible alternatives as possible ("forking paths"). Choices are considered defensible if they are reported in the literature and are feasible. Although this approach is necessarily less objective than basing decisions on a systematic review of the literature, the inclusion of the opinions of researchers from various labs likely leads to a relatively representative selection of alternatives. Following identification of all defensible processing paths, to facilitate the interpretation of the results, all members vote on a "preferred path", based on theoretical or practical preferences of each lab, which is then preregistered and used to address the hypothesis

in question. Thus, rather than basing analysis decisions on a potentially idiosyncratic path preferred by one researcher, conclusions are based on analysis decisions preferred by the majority of a team of experts. This also greatly reduces uncertainty in cases where individual researchers may not yet have developed an idiosyncratic routine for certain analysis steps. Finally, in order to disclose flexibility in data analysis, a Forking Path Analysis is applied for every tested hypothesis, across a random subsample of at least 1000 of all identified pre-processing, quantification and statistical choices ("forking paths"). This will demonstrate the relative frequency of defensible pre-processing and analysis paths leading to the same statistical finding as reported by the preferred path. With this approach, we attempt to assess the robustness of the preferred path and describe the variance related to existing alternative analysis routines. In a future step, the results of these Forking Path Analyses applied across different experimental settings and EEG components can help to validate, improve, or formulate practical recommendations and standardized pipelines.

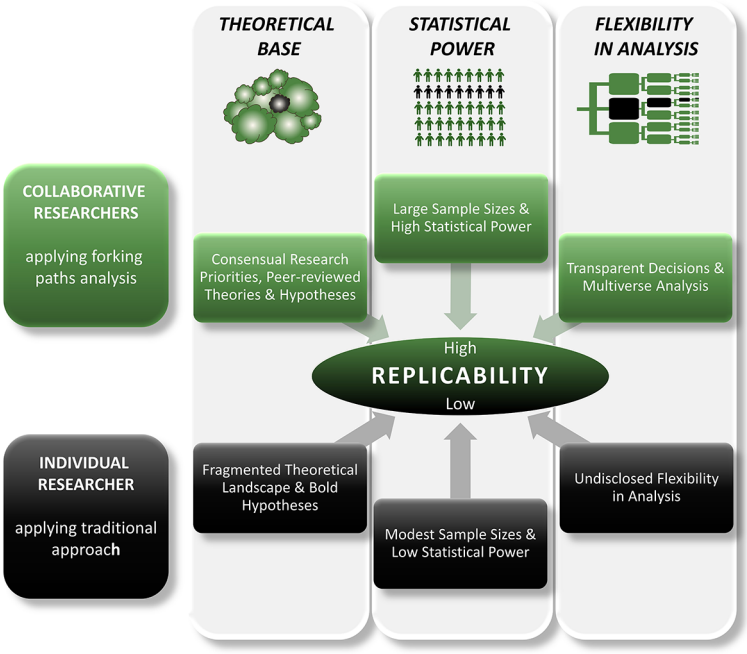
Our application of the cFPA approach complements existing projects aimed to increase open science practices and replicability such as the EEGManyLabs and EEGManyPipelines initiatives. However, it differs in two important aspects. On the one hand, the research questions of interest address inter-individual differences, where the effect sizes are known to be small ($r < .3$). In a classical EEGManyLabs project, the approach of conducting a meta-analysis of the results obtained in each participating lab would require each lab to collect an unrealistic amount of data. On the other hand, and in contrast to the EEGManyPipelines project, we started with a group of members already aware of the processing and analysis variance, and maintained a central component of cooperative collaboration throughout the entire project from planning to publication. Instead of "just identifying" this variance, we also aim to work towards "consensus" building, which aids the drawn conclusions. Simultaneously, we use the data to verify how robust the agreed path is relative to other defensible paths and identify variance associated with different analytical choices. Additionally, we plan to compare our cooperatively agreed path with more lab-specific pipelines. Although we embrace the EEGManyLabs and EEGManyPipelines initiatives, we feel that data should not only be combined in a technical or meta-analytical way, but that proactive consensus-building across research teams regarding the conception of hypotheses and data analyses is also beneficial. In this sense, the current CoScience project can be regarded as a novel way of stimulating scientific exchange beyond sharing of data, results, and data-analytic codes.

Fourth, it is still *uncommon for researchers to share their datasets* with other researchers. Given the organizational efforts, running times, and additional funding requirements associated with organizing large-scale projects, the availability of resulting datasets for additional research questions is essential. The cFPA approach supports the sharing of the large and diverse resulting dataset and analysis scripts for use by fellow researchers.

A visual representation of the differences between the research process in a single-laboratory and the cFPA research process and their impact on replicability is illustrated in Figure 1. It should be noted that typical differences have been presented and highlighted for illustrative purposes. We fully acknowledge ongoing efforts to reduce these problems in single-laboratory research, e.g., by preregistration and increasing sample size. Furthermore, we do of course not mean to imply that collaborative approaches should completely replace single-laboratory research. For instance, an important advantage of single-laboratory research is that no broad consensus across researchers is necessary, facilitating new, possibly unusual and creative hypotheses and methods.

**Figure 1**

*An Illustration of the Common Issues That Impact Replicability in Different Ways*




*Note.* The lower pane includes tendencies that can occur more easily in individual teams, while the upper pane shows how a more collaborative approach could help to tackle these issues.

Nonetheless, we are convinced that the collaborative nature of the CoScience EEG-personality project enables several otherwise unobtainable benefits. This article provides a detailed description of the initial application of this approach including project design and experimental procedure. The in-depth methodological description serves to enhance consistency in the methodology reported across all empirical articles resulting from this

comprehensive dataset, and aims to increase its visibility to potential future collaborators. Finally, we will inform readers of the process by which future collaboration with this dataset is possible.

# Project Preparation

This section describes the implementation of cFPA within the project. It is structured using three of the five principles of cFPA as laid out by Wacker (2017); namely cooperation between collaborators, the agreed-upon design, and the multicenter distribution of the data collection.

## Cooperation

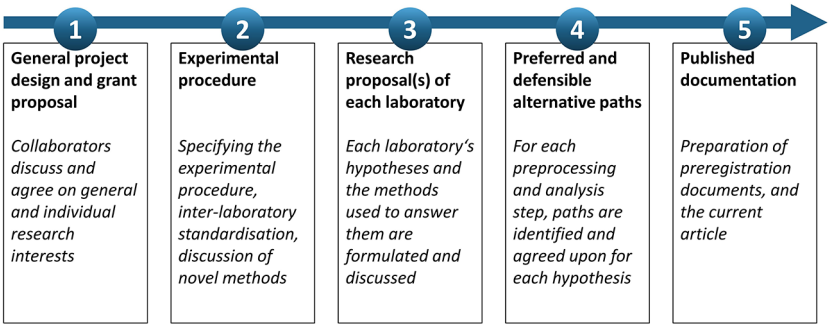
Ten principal investigators (PIs) actively contributed to the project funded by the German Research Foundation (DFG), nine of whom were co-applicants for the grant (DFG project number: 409321828). All PIs are scientists studying individual differences in Personality Neuroscience. Each has published in EEG-Personality Neuroscience and has the means to contribute to the EEG data collection effort. In addition to the PIs, two post-doctoral researchers (the two first authors of this article) were fully dedicated to the organization of the project. Several junior members of the PIs' groups were actively involved.

## Agreed-Upon Design

A fundamental principle of the project is that all components of the design and experimental procedure are agreed-upon in open, systematic scientific discourse. This includes the experimental procedure, research questions, and defensible processing and analysis paths. This was achieved through a series of stages, using an online forum and in-person and online workshops to communicate openly, rigorously, and democratically. Each stage of the process is described below in chronological order. A schematic illustration is provided in Figure 2.

**Figure 2**

*Schematic Illustration of the Chronological Process Completed to Achieve All Agreed-Upon Elements of the Project*




(1) Nine experts in EEG-Personality research contributed to a successful joint grant proposal. This included each laboratory's research questions and suggestions on the methods required to answer them. (2) All involved PIs and several junior members agreed upon the specific experimental procedure. This included the final data collection process, the within-task procedures, and the extent to which inter-laboratory standardization would be achieved. (3) Every research hypothesis was reviewed by collaborators using a customized online forum (www.coscience.net), which provided the ground rules of collaboration and made the discussion open and transparent to all project members. The dynamic open collaborator-review process consistently produced acceptance of the final research proposal. This enabled all research proposals to benefit from a wealth of expertise, encouraged methodological coordination between laboratories, and provided a basis for the preregistration documents. (4) To apply the forking path analysis, PIs and several junior members first identified all defensible processing and analysis decisions for each hypothesis and then voted on the preferred option. These discussions and decisions were based on previous literature, existing pre-processing pipelines, and collective expertise. This provided the forking and preferred paths for each hypothesis (restricting the statistics to general linear models). (5) All published documentation, including preregistration documents and resulting manuscripts, are collaborator-reviewed and approved within the group. Due to the intense collaboration, all PIs are listed as co-authors in the final publications. Junior members were invited to self-nominate to contribute to the publications.

## Multicenter Datasets

A total number of 720 participants with equal distribution of 72 per laboratory was agreed upon, which provides an appropriate sample size to test small associations and inter-lab variance. For instance, with this sample size per laboratory, the power is adequate

(i.e., .80) for a moderate association of $r$ = .28 with a one-tailed test (alpha = .05). Data collection was scheduled to be completed by all laboratories by the end of June 2022.

The equipment used during the laboratory-based sessions was standardized across laboratories on a general level, including experimental computers running a Windows operating system; a 24 inch stimulus display (Dell U 2412M) with a 60Hz refresh rate placed 25-35 inches from the participant's head; identical file scripts in PsychoPy (2020.1.3, Peirce, 2007); and an identical response box (MilliKey MH-5). All EEG recordings were completed using gel-filled Ag-AgCl electrodes according to the 10-20 system. All labs recorded mastoids and recorded horizontal and vertical ocular movements in unipolar, using between two and four electrodes. Electrocardiography (ECG) was recorded in bipolar mode by electrodes placed on the right wrist and 10cm above the left ankle, with a ground electrode placed on the right ankle based on Einthoven II placement in laboratories where this was required. Data were recorded with a sampling rate of 500 and 512 Hz for BrainProducts and Biosemi amplifier systems, respectively. All laboratories collected saliva samples using OG-500 Saliva Self Collection Kit from Steinbrenner Laborsysteme GmbH. Samples were frozen to enable subsequent analysis of genetic data depending on the availability of additional funding.

Despite these consistencies, some variance in the equipment features existed between laboratories. These inconsistencies are presented in Supplementary Table 1 and relate to the use of different recording systems, which affects sampling rate, filter settings, and electrode types. In order to control for these inconsistencies, inter-lab variance will be factored into the analyses by means of a multilevel model (by treating individuals as nested within laboratories and by including a random intercept and, if needed, a random slope across laboratories). This extension will be applied to the preferred path for testing the main hypothesis of each paper. Ultimately, the project can help to identify and describe possible sources of inter-laboratory variance across multiple tasks and EEG components.

## Forking Paths Analysis

Following acquisition of the multicenter dataset, a cFPA will be performed for each preregistered hypothesis. Each hypothesis-specific preferred path and defensible alternatives are included within each document detailing a preregistered hypothesis (see Supplementary Materials). To reduce the number of identified processing and analysis paths, we use only default parameters (e.g., for artefact correction algorithms) and distinctive parameters (e.g., for filtering). Nevertheless, with respect to pre-processing alone, we identified 14 relevant steps (from resampling, over filtering to artefact correction) involving 57 choices which, together, yielded 18 Million combinations (forking paths). Similarly, we identified around 1,500 ways on how to quantify an ERP component (including choices on which trials to include, baseline period, electrodes or time window). A

detailed description of all considered forking paths for an example research question can be found in the Supplementary Materials.

## Experimental Procedure

The present section provides a summary of the experimental procedure followed by each participant, categorized into three stages: recruitment, online assessment, and the laboratory session. A more detailed description and the experimental procedure (in PsychoPy) can be found in the Supplementary Materials.

### Recruitment

The study protocol was approved by the ethics committee of the German Society of Psychology (DGPs). Participants were recruited through online and offline postings at the university campus, university-related platforms, and social networks. Eligibility criteria, assessed through self-reports, were age of 18–30 years, heterosexuality (due to preregistered hypotheses regarding stimulus'/experimenters' attractiveness), right-handedness, fluency in German, ability to provide informed consent, and no current or past physical or psychological disease, no prescription of drugs, no drug abuse, no regular smoking, and no dreadlocks. Data collection comprised two stages, with an online administration of several self-report questionnaires and an in-laboratory session. Female participants were requested to schedule the laboratory session to fall within either 0–4, 7–11, 13–17, or 20–26 days since the first day of their last period. For completion, participants were reimbursed with 10€ per hour or course credits plus predetermined in-task winnings of 12.50€.

### Online Assessment

The online assessment was conducted using LimeSurvey. Upon giving informed consent, participants provided demographic information and completed a series of 18 personality questionnaires containing 14 intermittent attention check items (e.g., "I never used a computer."). Completion took approximately 1.5 hours. A list of the questionnaires is presented in Table 1 in order of administration. After completion of the online questionnaires and scheduling of the laboratory session, all participants received on-the-day instructions for the laboratory visit and a photograph of the experimenter to facilitate (and later test) interactions between participant and experimenter despite the mandatory facemask due to the ongoing COVID-19 pandemic.

**Table 1**

*Overview of All Collected Variables in This Project*

| **Assessed Variables in the COSCIENCE EEG-PERSONALITY PROJECT** | |
|---|---|
| **Demographic Variables** | |
| sex, age, ethnicity, relationship status, employment status, job title, highest academic qualification, average gross monthly income, use of hormonal contraceptives, handedness | |
| **Personality Variables** | |
| **Big Five** Inventory 2 (BFI-2; Danner et al., 2016) | **Intellect** Scale (Mussel, 2013) |
| **Life Events** (adapted from Brugha & Cragg, 1990) | **Behavioral Inhibition/Approach System** (BIS/BAS; Strobel et al., 2001) |
| Multidimensional **Perfectionism** Scale (MPS-F; Altstötter-Gleich & Bergemann, 2006) | **Anhedonia** Subscale of the Personality Inventory of the DSM-5 (PID-5; Zimmermann et al., 2014) |
| **Well-Being** Index (WHO-5; World Health Organization, 1998) | **Positive Emotionality** Scale of the Multidimensional Personality Questionnaire (MPQ BF; Angleitner & Ostendorf, 1993) |
| **Aggression** Questionnaire (Werner & von Collani, 2004) | Brief **Self-Control** Scale (BSCS; Sproesser et al., 2011) |
| **Reinforcement Sensitivity** Questionnaire (RST-PQ; Pugnaghi et al., 2018) | Brief **Childhood Trauma** Questionnaire (CTQ; Klinitzke et al., 2012) |
| **Fear of Negative Evaluation** Scale (SANB-5; Kemper et al., 2011) | Penn State **Worry** Questionnaire (PSWQ; Glöckner-Rist & Rist, 2014) |
| **Need for Cognition** (NFC; Bless et al., 1994) | **Prosocial Tendency** Measure (PTM-R; Rodrigues et al., 2017) |
| Beck **Depression** Inventory (BDI-II; Hautzinger et al., 2009) | |
| **Effortful Control** from the Adult Temperament Questionnaire (ATQ; Wiltink et al., 2006) | |
| **Experimental Task Variables** | |
| **Rest**: instruction (eyes open/closed), repeated three times throughout the experiment | |
| **Go-NoGo**: stimulus type (Go/NoGo), instruction (emphasis on relaxation or accuracy) | |
| **Gambling**: feedback valence (reward/loss), feedback magnitude (0, 10, 50 cents) | |
| **Emotional Stroop**: picture category (erotic couple, erotic woman, erotic man, positive, neutral, neutral couple, tree) | |
| **Flanker**: congruence of flanker and target (100, 66, 33, 0%), social evaluation (experimenter present/absent), two ratings on work-load | |
| **Ultimatum Game**: offer fairness (fair, moderately fair, unfair) | |
| **Additional Data** | |
| **Intelligence**: assessed through matrices (fluid) & knowledge tests (crystallized) (Intelligence-Structure-Test 2000 R; Liepmann et al., 2007) | |
| **Self-reported mood:** anxious, full of pep, peeved, happy, tired, relaxed, sad, exhausted, and irritated, collected nine times throughout the experiment | |
| **Physiological data:** 64-channel continuous recording of EEG, eye movements, mastoids, and ECG | |
| **Molecular genetic data:** saliva sample | |
| **Reciprocal participant and experimenter ratings:** Big Five, familiarity, attractiveness, competence | |
| **On-the-day participant variables:** consumption of food, coffee, drugs, sleep, cycle phase (female only) | |
| **Laboratory variation:** operating system, EEG amplifier model & EEG cap make (BioSemi, BrainProducts), electrodes (passive, active), ground and reference electrodes | |
| **Environmental and other variation:** room temperature, time of day, and exact onset of visual stimuli | |

*Note.* For details see the following sections of this manuscript or the online preregistration on OSF (osf.io/yq3z7).

## Laboratory Session

Participants completed a series of tasks within one session, which are listed in order in Figure 3, while EEG and ECG were recorded. Within-task procedures and timings are presented in Figure 4. Task descriptions are provided below.

**Figure 3**

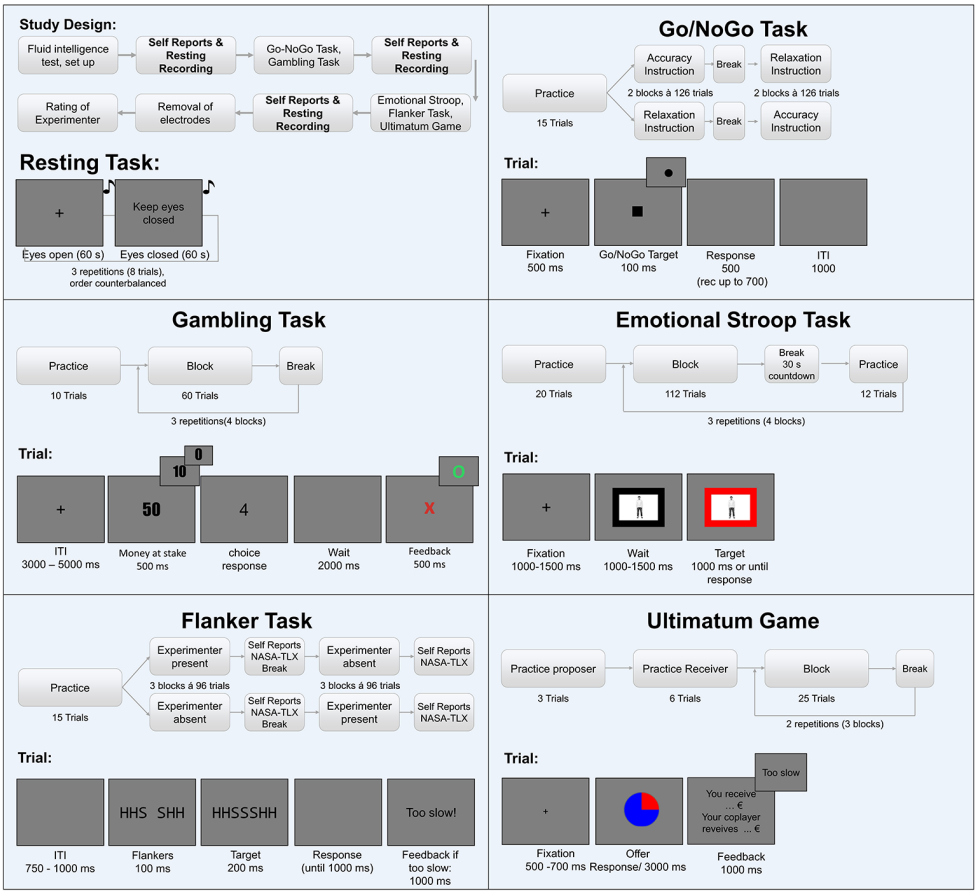
*A Schematic Illustration of the Overall Experimental Procedure of the Laboratory Session, Including Duration of Each Stage and the Total Duration*

| Task | Duration [min] | Time since start [h:min] |
|---|---|---|
| Arrival questions + fluid intelligence test | 12 | 0:12 |
| Application of EEG and ECG electrodes | 30 | 0:42 |
| Eye calibration | 2 | 0:44 |
| Self-report + EEG and ECG rest measurement | 12 | 0:56 |
| Go-NoGo Task + two self-reports | 37 | 1:33 |
| Gambling Task + self-report | 37 | 2:10 |
| EEG and ECG rest measurements | 8 | 2:18 |
| Break | 15 | 2:33 |
| Emotional Stroop Task + self-report | 18 | 2:51 |
| Flanker Task + three self-reports | 30 | 3:21 |
| Ultimatum Game + self-reports | 15 | 3:36 |
| EEG and ECG rest measurements | 8 | 3:44 |
| Removal of EEG and ECG electrodes | 20 | 4:04 |
| Rating of experimenter | 3 | 4:07 |
| Crystallized intelligence test | 35 | 4:42 |
| Saliva sample | 3 | 4:45 |
| Rating of participant | 3 | 4:48 |

*Note.* A blue background indicates that EEG and ECG data was recorded at this time. A white background indicates that no EEG or ECG was recorded at this time. min = minutes; h = hour(s).

**Figure 4**

*Task Designs of the Resting Measurement, Go-NoGo Task, Gambling Task, Emotional Stroop Task, Flanker Task, and Ultimatum Game*




## Arrival Questions

Upon arrival, participants indicated whether their coffee and/or food consumption that day, and sleep the night before deviated from typical for them and, if so, in which direction. Participants indicated how much alcohol they consumed within the previous 24 hours on a scale of 0, 1–3, 4+ units, confirmed whether or not they had eaten a large meal within the previous hour, and whether or not they had consumed other recreational drugs within the previous 24 hours. Participants also confirmed how many cups of coffee they had consumed that day, and how many hours they had slept the night before. In addition, females reported the number of days since the first day of their last period.

### Brief Fluid Intelligence Test

Twenty matrices items assessing fluid intelligence were selected from the Intelligence-Structure-Test Revised (I-S-T 2000 R; Liepmann et al., 2007). The tasks were provided to participants on paper, and answers were selected on the computer using the mouse. They had 10 minutes to finish the test.

### Crystallized Intelligence Test

Eighty-four items assessing knowledge were selected from the I-S-T 2000 R (Liepmann et al., 2007). The tasks were provided to participants on paper, and answers were selected on the computer using the mouse. They had 40 minutes to finish the test.

### Self-Reports

To capture mood changes throughout the experiment, a series of ten self-report mood questions were completed nine times during the laboratory session, spaced intermittently between and within tasks as illustrated in Figure 3. The items were presented in a fixed order. Similar to the Profile of Mood States (POMS, Albani et al., 2005; after McNair et al., 1992), participants indicated the extent to which each mood reflected their present mood state using a 7-point Likert scale ranging from *überhaupt nicht* (not at all) to *sehr stark* (very strong). In order of presentation, the items were *ängstlich* (anxious), *schwungvoll* (full of pep), *verärgert* (peeved), *fröhlich* (happy), *müde* (tired), *entspannt* (relaxed), *betrübt* (sad), *ruhig* (calm), *erschöpft* (exhausted), and *gereizt* (irritated).

### Go-NoGo Task

A response inhibition Go-NoGo Task based on the paradigm introduced by Verbruggen and Logan (2008) was used. Within this paradigm, participants respond to visual cues (*Go* trials) while withholding response to others (*NoGo* trials). Squares and circles were used as *Go* and *NoGo* stimuli, counterbalanced across participants. *Go* trials were presented more frequently (66%) than *NoGo* trials, and the trial order was randomized within the two experimental blocks. The experimental blocks differ in their instructions, the order counterbalanced between participants. In a string of two consecutive blocks, participants were instructed to perform as fast and as accurately as possible. In the other two consecutive blocks, participants were instructed to be as relaxed as possible, while responding accurately, and were reassured that relaxation is more important than correct performance. Participants answered the self-report mood questions between the two blocks.

### Gambling Task

Based on a paradigm used by Sato et al. (2005), participants guessed whether an undisclosed number would be higher or lower than a disclosed number, to win or lose a varying amount of money depending on whether they guessed correctly or incorrectly,

respectively. Participants were informed that they would receive their winnings at the end of the experiment. Within each trial, participants were first informed of the magnitude of the potential win or loss, which varied between 0, 10, or 50 cents. Following this, participants were shown the disclosed number and asked to indicate whether an undisclosed number would be higher or lower than the disclosed number. Afterwards, win or loss feedback was presented. Unbeknown to the participant, the feedback was pre-programed, independent of response, and win and loss feedback was presented equally often. Participants were shown the numbers 4 and 5 most frequently, wherein both outcomes (i.e., a higher or lower undisclosed number) were similarly likely. Remaining numbers (2, 3, 6, 7, 8) were shown in 16% of the trials. The disclosed numbers, magnitude of win/loss, and feedback were randomized across trials.

### Emotional Stroop Task

As an adaptation of the Emotional Stroop paradigm used by Munk et al. (2020), participants were presented with a series of individual images, first with a black frame, which changed into a differently colored frame. Participants responded to the color of the frame by pressing the corresponding colored button on the response box. There were four colors of frame: red, green, blue, and yellow. There were seven image categories: erotic couple, erotic woman, erotic man, positive (man or woman), neutral, neutral couple, and tree. Each category contained eight different images. Each image was shown eight times (twice with each colored frame), with image-frame combinations presented randomly across blocks. After the task, participants provided one rating of affective valence and one rating of arousal for each category of images using the self-assessment-manikin and the corresponding number on the keyboard (Bradley & Lang 1994).

### Flanker Task

As in Forster et al. (2011), participants responded with their right hand to a central letter (target) while ignoring the letters on either side of the target (flankers). Written feedback informing the participant they were too slow (*zu langsam*) was presented in response to reactions slower than 1000 ms. Response mappings were counterbalanced between subjects. Task demand varied between trials due to variance in the congruence level of the flanker stimuli. In half of the trials, flanker stimuli were 100% congruent (e.g., SSSSSSS), in the other half they were either 67% congruent (e.g., HSSSSSH), 33% congruent (e.g., HHSSSHH), or 0% congruent (e.g., HHHSHHH). Trials were presented in a randomized order across blocks of similar social pressure. Social evaluation was manipulated across blocks. During one string of three blocks, the experimenter was present in the room. During the other string of three blocks, the experimenter was absent from the room. The order of the experimenter's presence and absence was randomized between participants. At the end of the experimental blocks, participants filled in an

adaptation of the task load index (NASA-TLX) and the self-reported mood state in the format described above.

### Ultimatum Game

This version of the Ultimatum Game was adapted from Rodrigues et al. (2015). Participants first acted as proposers to become acquainted with the procedure. In the experimental phase participants acted as responders, receiving offers supposedly made by other participants, and choosing either to accept and receive their proportion of the money, or reject so nobody received any money. Offers varied in their fairness level (9:1, i.e., proposer received 9 cents while the responder received 1 cent, 5:5 or 7:3). Each offer was presented 25 times in a randomized order across three blocks.

### Rest Measurements of EEG and ECG

Three rest measurement sessions occurred throughout the experimental session, as indicated in Figure 3. Each session comprised eight minutes of EEG and ECG rest recordings, within which single minutes of eyes-open and eyes-closed alternated. A 500 ms tone provided a signal to the participant to change between these states.

### Rating of the Experimenter/Participant

At the end of the experiment, participants rated the experimenter's personality (using BFI-10, Rammstedt et al., 2012), attractiveness, likeability, dominance, and competence, in a fixed order. The latter four measures comprised a single item each and were structured using the same answer format as the BFI-2. Finally, participants indicated how familiar they were with the experimenter. After each participant left, the experimenter rated the participant on the same items.

### Saliva Sample

Each participant provided a single, self-collected saliva sample at the end of the data collection process using self-collection saliva-sample kits. Participants were provided with a 5-step illustrated sheet of instructions on how to provide and secure an uncontaminated sample correctly.

## Applications of This New Dataset

The current dataset includes a large number of high-quality EEG, ECG, and behavioral measures as well as a wide range of self-reported personality traits, conducive to answering a plethora of (EEG-personality) research questions. Further, we include suggestions for future endeavors, where we welcome additional experts in the field. Finally, we outline the availability of this new dataset.

## Research Avenues

### Examples of Existing Preregistered Research Questions

Various hypotheses have been preregistered by project members. A brief overview of a small subset of these is described below, whereas specific and detailed individual hypotheses can be found in the Supplementary Materials.

**Frontal EEG Asymmetry and Personality** — For over 20 years, left versus right frontal "resting" cortical activity (as inferred right versus left inhibitory EEG alpha activity) has been discussed as a marker of trait-like individual differences in approach and withdrawal motivation (Davidson,1998). Although by now, dozens of independent studies have investigated the link between this EEG measure and various traits, a meta-analysis (Kuper, Käckenmester, & Wacker, 2019) revealed that no reliable association has emerged. At the same time, studies treating EEG asymmetry as a state variable indexing the degree of current approach or withdrawal motivation have yielded more robust associations. Therefore, we are going to test the effects of (1) motivational factors during different experimental task settings (such as the Gambling and the Emotional Stroop task), as well as (2) typically unreported situational variables (such as experimenter attractiveness). It is hypothesized that associations between frontal alpha asymmetry and traits are strongest in motivationally relevant situations (e.g., (1) when a lot of money is at stake vs. not, (2) when experimenter is perceived as attractive vs. not).

**Error- and Feedback-Related Negativities (ERN and FRN) and Personality** — Both ERN and FRN are negative deflections in the event related potential (ERP) at frontocentral scalp locations with the former occurring 50 to 150 ms after execution of an error response and the latter 200-300 ms after a worse-than-expected feedback. While meta-analytical evidence suggests a small association between these components and anxiety/neuroticism and depression (Cavanagh & Shackman, 2015; Moser et al., 2013), studies vary largely in the applied trait measures and reported effects. Moreover, the strength of these associations varies greatly across individuals and situational contexts, as differences were reported for men and women, as well as for situations impending a threat or stressor (Osinsky et al., 2017). Therefore, we are going to (1) test the specificity of the associations of the ERN/FRN with various traits related to anxiety and depression, and (2) test the effects of situational variables on this relationship. As such, it is hypothesized that (1) the association between the FRN (in a Gambling task) and depression is driven by anhedonia and low reward sensitivity, and (2) that the association between the ERN (in a Flanker task) and dispositional anxiety is stronger when a social stressor (observing experimenter) is present.

**Quantifying Neural Activity in an EEG-Personality Approach** — An often-neglected source of (error) variance in ERP studies originates from the chosen index of the ERP

amplitude (e.g., peak amplitude, peak-to-peak amplitude, mean amplitude) or reference method (Common Average, Mastoids, CSD). At the same time, new methodological advances such as the Gamma Model Approach (Stahl et al., 2010) make it possible to capture individual variations in ERP parameters in great detail (width, rise, onset). However, many of these decisions or new approaches have not been tested in a large dataset. Capitalizing on this new dataset, we will (1) systematically compare quantification choices of the FRN amplitude (during the Ultimatum game) as well as (2) validate the gamma model analysis in the context of the ERN (during a Flanker and Go-NoGo task). We hypothesize that (1) quantification choices impact reliability of the measured ERP components and their association with personality traits, and (2) several parameters of the Gamma Model Approach (scaling, inflection point, skewness…) can differentiate correct and incorrect actions.

### Avenues for Future Approaches

The complexity of the multi-dimensional dataset and the innumerable ways to analyze it make it impossible to provide a complete list of potential dependent variables or analysis options. This following list functions to inspire new perspectives on the dataset and to present additional research opportunities.

**Structure of Personality** — The current dataset includes a variety of different personality questionnaires, assessing traits that are, by definition, similar. By considering these self-reports with behavior and neurophysiological indices, the relationship amongst these traits can be revisited and tested.

**Temporal Changes** — Several measures are recorded several times over the experiment such as mood, EEG and ECG data at rest. Therefore, this dataset enables the investigation of the impact of the experimental procedure on these state variables, or how these state variables affect cognitive control processes such as error monitoring.

**Multimodal Assessment** — In addition to EEG, ECG data is recorded during the experiment. While we preregistered one hypothesis concerning cortico-cardiac coupling during the gambling task, this measure can reveal insights into activity of the vegetative nervous system, in both task-related fashion (heart rate acceleration/deceleration to certain events) as well as in a tonic fashion (such as heart rate variability changes at rest or during specific task blocks).

**Advanced Analysis Methods** — Until now, the preregistrations focused mainly on classical components of the EEG signal in the time-domain (such as ERP components like ERN, FRN, N2, LPP) and the frequency-domain (frontal alpha asymmetry, mid-frontal theta). These are just a fraction of EEG components, which could be studied in terms of their reliability or their associations to inter-individual differences. Other markers may

include ERPs such as the SPN, P3, CNV, N1, P2, or power in other frequency domains such as beta or delta. However, this large dataset allows to venture out and apply other quantifications of the EEG signal such as single-trial analysis, topographical analysis, source estimations, coherence measures, or network estimates.

**DNA Analysis** — Saliva samples were collected to provide material for genetic analyses pending additional funding. We plan to use micro-arrays for DNA analysis allowing us to derive polygenetic scores and many other aggregate genetic indicators that may be used fruitfully with the large sample size.

**Methodological Investigations** — This project offers the unique opportunity to gain access to scripted automatic pre-processing pipelines for multiple defensible alternative EEG pre-processing sequences across various EEG components. This is highly beneficial to researchers wishing to investigate the influence of processing and analysis decisions on hypothesis-specific outcomes or on EEG component data more generally, to further inform methodological guidelines.

**Reliability** — The reliability of different methods of ERP quantification are not regularly compared within the same data set, even though this may guide decisions implicitly. Therefore, specification of analytic approaches as *a priori* has been recommended (Klawohn et al., 2020). The large data set provided by the present project allows systematic comparison of the reliability of different ERP quantifications, in addition to different methods for the estimation of ERP reliabilities across subsamples.

**Machine-Learning** — The high-dimensionality of this dataset is conducive to a high potential for machine learning to model complex relationships within and between personality, cognitive, and physiological variables. In addition to the high-dimensionality, the large sample size offers the opportunity for machine learning to apply resampling methods to assess the generalizability of these models.

## Data Availability for Researchers

In the spirit of fostering open science and transparency, researchers around the world may gain access to the full dataset to answer their own research questions (explicitly including reanalyzes for previously addressed hypotheses). However, since the multi-dimensional data contains potentially identifiable information (in particular but not exclusively with the genetic material), it cannot be made publicly available through common means such as the Open Science Framework. Instead, researchers are required to contact the project lead. To gain access to the project's data at no cost, data requesters must join the online forum for this project (www.coscience.net). The forum is closed to accepted members in order to ensure that the open collaborator-review process is carried out in

a fair and supportive environment. Membership is granted to researchers who adhere to the *Code of Conduct* (https://coscience.net/pdf/coscience_code-of-conduct.pdf) and the *Terms and Conditions* (https://coscience.net/pdf/coscience_terms-and-conditions.pdf), which strive to increase transparency, open communication, high levels of scientific rigor, and mutual support. Upon approval of membership, new members submit their data request to the forum using the templates provided. The data request includes a detailed description of the background, hypothesis, and analysis plan. This is followed by an open collaborator-review by the existing members of the project. The collaborator-review process is carried out to make sure that hypotheses are scientifically sound, sufficiently distinct from existing preregistrations, and ethically justifiable. Moreover, the data requester can benefit from the provided feedback to sharpen hypotheses or improve analytical decisions. However, collaborative-reviewers do not act as co-authors unless specifically agreed upon by the data requester.

This open and simple process is beneficial to future collaborators. Through this internal collaborator-review process, we aim to sharpen hypotheses, improve analytical choices, minimize overlap between different data requests, and ensure completeness of the resulting preregistration. Following this process, data requesters receive a permuted random subsample of raw data (with each variable or set of variables shuffled independently) from which they are expected to create a complete analysis script based on Matlab and R code up to the final statistical test(s) of interest within six months. Data requesters will likely benefit from the wide range of scripts already prepared within the project. The new script will then be applied to the original complete data set by the lead investigator's team in collaboration with the data requester. After receiving all analysis results, data requesters are expected to submit an initial manuscript draft to be reviewed by all group members who actively contributed to the manuscript. Data requesters who take the lead for further optimization of the manuscript after (re)submission to an international peer-reviewed journal will serve as first authors, and only CoScience team members who contribute are considered as co-authors.

## Caveats and Limitations

Implementation of this large-scale project was not without limitations. First, the collaborative and democratic approach required high commitment from all participating laboratories and orchestration by the leading laboratory. This resulted in an extensive planning phase of high organizational effort wherein previous decisions were revisited as new evidence was gathered. Second, the presence of the COVID-19 pandemic delayed the start of data collection greatly and may have influenced the final sample. The pandemic and accompanying additional requirements, such as wearing a mask and obtaining a negative test result, may have enhanced the barrier for participation for highly anxious individuals. However, the comprehensive array of personality variables recorded will allow identification of such bias. Third, the quality of the agreed upon preferred analysis

path is limited to the shared expertise of the collaborators. However, this would be limited further to the expertise of one in a single-laboratory paradigm. Forth, the inclusion of multiple laboratory paradigms required participants to complete a five-hour laboratory session, which may have influenced engagement in later stages of the experiment. However, a mid-session break wherein participants were provided with refreshments, helped to mitigate a reduction in engagement. Additionally, the intermittent completion of self-reported mood scales throughout the session could be analysed as a by-proxy measure of engagement. Because the focus of the project is on individual differences, we opted against randomizing the order of the tasks, as this would introduce an additional source of between-subject variance that may conceal the small to moderate effects of interest even with the study's large sample size. Instead, we implemented small breaks in between each task that ended once the participant indicated they felt ready.

## Conclusion

The present article outlines a large multicenter dataset yielded from a large EEG-personality project grounded in the principles of cFPA. The dataset alleviates the reproducibility issues of low power and researcher degrees of freedom, and is appropriate for answering a plethora of EEG-personality research questions. This article should be cited in all empirical publications that are based on the aforementioned dataset, to provide depth to the reported methodology. Project members look forward to welcoming future collaborators to the project.

**Funding:** A research grant from the Deutsche Forschungsgemeinschaft (DFG) covered the costs associated with the realization of the CoScience project described in the manuscript. Grant numbers: WA2593/9-1, MU3535/7-1, BE 2443/12-1, STA 1035/6-1, HI 1780/6-1, STR 615/7-1, HE 5330/15-1, HE2443/16-1, OS422/7-1. The funders had no role in study design, data collection, data analysis, or preparation of the manuscript. There was no additional external funding received for this project.

**Acknowledgments:** The authors have no additional (i.e., non-financial) support to report.

**Competing Interests:** The authors have declared that no competing interests exist.

**Author Contributions:** *Katharina Paul*—Idea, conceptualization | Design planning | Visualization (data presentation, figures, etc.) | Validation, reproduction, checking | Writing | Feedback, revisions | Project coordination, administration. *Cassie Ann Short*—Idea, conceptualization | Design planning | Visualization (data presentation, figures, etc.) | Validation, reproduction, checking | Writing | Feedback, revisions | Project coordination, administration. *André Beauducel*—Design planning | Validation, reproduction, checking | Feedback, revisions | Funding to conduct the work. *Hannes Per Carsten*—Design planning | Validation, reproduction, checking | Feedback, revisions. *Kai Härpfer*—Design planning | Validation, reproduction, checking | Feedback, revisions. *Jürgen Hennig*—Design planning | Validation, reproduction, checking | Feedback, revisions | Funding to conduct the work. *Johannes Hewig*—Design planning | Validation, reproduction, checking | Feedback, revisions | Funding to conduct the work. *Andrea Hildebrandt*—Design planning | Validation, reproduction, checking | Feedback, revisions | Funding to conduct the work. *Corinna Kührt*—Design planning | Validation, reproduction, checking | Feedback, revisions. *Erik Malte Mueller*—Design planning | Validation, reproduction, checking | Feedback, revisions | Funding to conduct the work. *Aisha Munk*—Design planning. *Roman Osinsky*—Design planning | Validation, reproduction, checking | Feedback, revisions | Funding to conduct the work. *Elisa Porth*—Design planning | Validation, reproduction, checking | Feedback, revisions. *Anja Riesel*—Design planning | Validation, reproduction, checking | Feedback, revisions. *Johannes Rodrigues*—Design planning | Validation, reproduction, checking | Feedback, revisions. *Christoph Scheffel*—Design planning | Validation, reproduction, checking | Feedback, revisions. *Jutta Stahl*—Design planning | Validation, reproduction, checking | Feedback, revisions | Funding to conduct the work. *Alexander Strobel*—Design planning | Validation, reproduction, checking | Feedback, revisions | Funding to conduct the work. *Jan Wacker*—Idea, conceptualization | Design planning | Visualization (data presentation, figures, etc.) | Validation, reproduction, checking | Feedback, revisions | Project coordination, administration | Funding to conduct the work.

**Author Note:** Katharina Paul and Cassie Ann Short are the joint first authors of the manuscript.

**Ethics Approval:** The project procedure was approved by the ethics committee of the German Society of Psychology (DGPs).

**Data Availability:** Since the multi-dimensional data contains potentially identifiable information (in particular but not exclusively with the genetic material), it cannot be made publicly available through common means such as the Open Science Framework. Instead, researchers are required to contact the project lead. To gain access to the project's data at no cost, data requesters must join the online forum for this project (www.coscience.net). The forum is closed to accepted members in order to ensure that the open collaborator-review process is carried out in a fair and supportive environment.

## Supplementary Materials

For this article the following Supplementary Materials are available (for access see Index of Supplementary Materials below):
Via the Open Science Framework (OSF) repository:

- Pre-registration
- Study design
- Experimental tasks (Psychopy files) presented in the CoScience Project

Via the PsychArchives repository:

- Table showing between-laboratory variation in the equipment used to collect data
- Open Peer Review

### Index of Supplementary Materials

Paul, K., Short, C. A., Beauducel, A., Carsten, H. P., Härpfer, K., Hennig, J., Hewig, J., Hildebrandt, A., Kührt, C., Mueller, E. M., Munk, A., Osinsky, R., Porth, E., Riesel, A., Rodrigues, J., Scheffel, C., Stahl, J., Strobel, A., & Wacker, J. (2021). *Supplementary materials to "The methodology and dataset of the coscience EEG-personality project – a large-scale, multi-laboratory project grounded in cooperative forking paths analysis"* [Pre-registration]. OSF. https://osf.io/g6u89

Paul, K., Short, C. A., Beauducel, A., Carsten, H. P., Härpfer, K., Hennig, J., Hewig, J., Hildebrandt, A., Kührt, C., Mueller, E. M., Munk, A., Osinsky, R., Porth, E., Riesel, A., Rodrigues, J., Scheffel, C., Stahl, J., Strobel, A., & Wacker, J. (2021). *Supplementary materials to "The methodology and dataset of the coscience EEG-personality project – a large-scale, multi-laboratory project grounded in cooperative forking paths analysis"* [Study design]. OSF. https://osf.io/yq3z7

Paul, K., Short, C. A., Beauducel, A., Carsten, H. P., Härpfer, K., Hennig, J., Hewig, J., Hildebrandt, A., Kührt, C., Mueller, E. M., Munk, A., Osinsky, R., Porth, E., Riesel, A., Rodrigues, J., Scheffel, C., Stahl, J., Strobel, A., & Wacker, J. (2021). *Supplementary materials to "The methodology and dataset of the coscience EEG-personality project – a large-scale, multi-laboratory project grounded in cooperative forking paths analysis"* [Experimental tasks]. OSF. https://osf.io/6b7ke

Paul, K., Short, C. A., Beauducel, A., Carsten, H. P., Härpfer, K., Hennig, J., Hewig, J., Hildebrandt, A., Kührt, C., Mueller, E. M., Munk, A., Osinsky, R., Porth, E., Riesel, A., Rodrigues, J., Scheffel, C., Stahl, J., Strobel, A., & Wacker, J. (2022). *Supplementary materials to "The methodology and dataset of the coscience EEG-personality project – a large-scale, multi-laboratory project grounded in cooperative forking paths analysis"* [Table]. PsychOpen GOLD. https://doi.org/10.23668/psycharchives.6671

Personality Science. (Ed.). (2021). *Supplementary materials to "The methodology and dataset of the coscience EEG-personality project – a large-scale, multi-laboratory project grounded in cooperative forking paths analysis"* [Open Peer-Review]. PsychOpen GOLD. https://doi.org/10.23668/psycharchives.6672

## References

Albani, C., Blaser, G., Geyer, M., Schmutzer, G., Brähler, E., Bailer, H., & Grulke, N. (2005). Überprüfung der Gütekriterien der deutschen Kurzform des Fragebogens „Profile of Mood States" (POMS) in einer repräsentativen Bevölkerungsstichprobe. *PPmP-Psychotherapie· Psychosomatik Medizinische Psychologie, 55*(07), 324–330. https://doi.org/10.1055/s-2004-834727

Altstötter-Gleich, C., & Bergemann, N. (2006). Testgüte einer deutschsprachigen Version der mehrdimensionalen Perfektionismus Skala von Frost, Marten, Lahart und Rosenblate (mps-f). *Diagnostica, 52*(3), 105–118. https://doi.org/10.1026/0012-1924.52.3.105

Angleitner, A., & Ostendorf, F. (1993). Zur Struktur von Fähigkeits-und Begabungsbegriffen in Selbst-und Bekanntenbeurteilungen. In H. Bauersfeld & R. Bromme (Eds.), *Bildung und Aufklärung: Studien zur Rationalität des Lehrens und Lernens: Festschrift für Helmut Skowronek zum 60. Geburtstag* (pp. 1–32). Münster, Germany: Waxmann.

Bless, H., Wänke, M., Bohner, G., Fellhauer, R. F., & Schwarz, N. (1994). Need for cognition: Eine Skala zur Erfassung von Engagement und Freude bei Denkaufgaben [Need for cognition: A scale measuring engagement and happiness in cognitive tasks]. *Zeitschrift für Sozialpsychologie, 25*, 147–154.

Botvinik-Nezer, R., Holzmeister, F., Camerer, C. F., Dreber, A., Huber, J., Johannesson, M., Kirchler, M., Iwanir, R., Mumford, J., Adcock, R. A., Avesani, P., Baczkowski, B. M., Bajracharya, A., Bakst, L., Ball, S., Barilari, M., Bault, N., Beaton, D., Beitner, J., … Schonberg, T. (2020). Variability in the analysis of a single neuroimaging dataset by many teams. *Nature, 582*(7810), 84–88. https://doi.org/10.1038/s41586-020-2314-9

Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry, 25*(1), 49–59. https://doi.org/10.1016/0005-7916(94)90063-9

Brugha, T. S., & Cragg, D. (1990). The list of threatening experiences: The reliability and validity of a brief life events questionnaire. *Acta Psychiatrica Scandinavica, 82*(1), 77–81. https://doi.org/10.1111/j.1600-0447.1990.tb01360.x

Button, K. S., Ioannidis, J. P., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S., & Munafò, M. R. (2013). Power failure: Why small sample size undermines the reliability of neuroscience. *Nature Reviews. Neuroscience, 14*(5), 365–376. https://doi.org/10.1038/nrn3475

Cavanagh, J. F., & Shackman, A. J. (2015). Frontal midline theta reflects anxiety and cognitive control: Meta-analytic evidence. *Journal of Physiology, Paris, 109*(1-3), 3–15. https://doi.org/10.1016/j.jphysparis.2014.04.003

Danner, D., Rammstedt, B., Bluemke, M., Treiber, L., Berres, S., Soto, C. J., & John, O. P. (2016). Die deutsche Version des Big Five Inventory 2 (BFI-2). *Zusammenstellung sozialwissenschaftlicher Items und Skalen (ZIS)*. https://doi.org/10.6102/zis247

Davidson, R. J. (1998). Anterior electrophysiological asymmetries, emotion, and depression: Conceptual and methodological conundrums. *Psychophysiology, 35*(5), 607–614. https://doi.org/10.1017/S0048577298000134

Forster, S. E., Carter, C. S., Cohen, J. D., & Cho, R. Y. (2011). Parametric manipulation of the conflict signal and control state adaptation. *Journal of Cognitive Neuroscience, 23*(4), 923–935. https://doi.org/10.1162/jocn.2010.21458

Gabard-Durnam, L. J., Mendez Leal, A. S., Wilkinson, C. L., & Levin, A. R. (2018). The Harvard Automated Processing Pipeline for Electroencephalography (HAPPE): Standardized processing software for developmental and high-artifact data. *Frontiers in Neuroscience, 12*, Article 97. https://doi.org/10.3389/fnins.2018.00097

Gignac, G. E., & Szodorai, E. T. (2016). Effect size guidelines for individual differences researchers. *Personality and Individual Differences, 102*, 74–78. https://doi.org/10.1016/j.paid.2016.06.069

Glöckner-Rist, A., & Rist, F. (2014). Deutsche Version des Penn State Worry Questionnaire (PSWQ-d). *Zusammenstellung sozialwissenschaftlicher Items und Skalen (ZIS).* https://doi.org/10.6102/zis219

Hautzinger, M., Keller, F., Kühner, C., & Beck, A. T. (2009). *Beck Depressions-Inventar: BDI II; Manual.* Frankfurt, Germany: Pearson Assessment.

Kemper, C. J., Lutz, J., & Neuser, J. (2011). SANB-5: Skala Angst vor negativer Bewertung-5 (SANB-5). In Leibniz-Zentrum für Psychologische Information und Dokumentation (ZPID) (Ed.), *PSYNDEX Tests. Datenbanksegment Psychologischer und Pädagogischer Testverfahren* [Doc. Nr.: 9006445]. Trier, Germany: ZPID. https://doi.org/10.23668/psycharchives.416

Klawohn, J., Meyer, A., Weinberg, A., & Hajcak, G. (2020). Methodological choices in event-related potential (ERP) research and their impact on internal consistency reliability and individual differences: An examination of the error-related negativity (ERN) and anxiety. *Journal of Abnormal Psychology, 129*(1), 29–37. https://doi.org/10.1037/abn0000458

Klinitzke, G., Romppel, M., Häuser, W., Brähler, E., & Glaesmer, H. (2012). Die deutsche Version des Childhood Trauma Questionnaire (CTQ): Psychometrische Eigenschaften in einer bevölkerungsrepräsentativen Stichprobe. *PPmP-Psychotherapie· Psychosomatik· Medizinische Psychologie, 62*(2), 47–51. https://doi.org/10.1055/s-0031-1295495

Kuper, N., Käckenmester, W., & Wacker, J. (2019). Resting frontal EEG asymmetry and personality traits: A meta–analysis. *European Journal of Personality, 33*(2), 154–175. https://doi.org/10.1002/per.2197

Liepmann, D., Beauducel, A., Brocke, B., & Amthauer, R. (2007). *Intelligenz-Struktur-Test 2000 R (erweiterte Auflage)* [Intelligence-Structure-Test; extended edition]. Göttingen, Germany: Hogrefe.

Luck, S. J., & Gaspelin, N. (2017). How to get statistically significant effects in any ERP experiment (and why you shouldn't). *Psychophysiology, 54*(1), 146–157. https://doi.org/10.1111/psyp.12639

McNair, D., Lorr, M., & Droppelman, L. (1992). *Profile of Mood States Manual (rev.).* San Diego, CA, USA: Educational and Industrial Testing Service.

Moser, J. S., Moran, T., Schroder, H., Donnellan, B., & Yeung, N. (2013). On the relationship between anxiety and error monitoring: A meta-analysis and conceptual framework. *Frontiers in Human Neuroscience, 7*, Article 466. https://doi.org/10.3389/fnhum.2013.00466

Munk, A. J., Dickhaeuser, L., Breitinger, E., Hermann, A., Strahler, J., Schmidt, N. M., & Hennig, J. (2020). Females' menstrual cycle and incentive salience: Insights on neural reaction towards erotic pictures and effects of gonadal hormones. *Comprehensive Psychoneuroendocrinology, 3*, Article 100006. https://doi.org/10.1016/j.cpnec.2020.100006

Mussel, P. (2013). Intellect: a theoretical framework for personality traits related to intellectual achievements. *Journal of Personality and Social Psychology, 104*(5), 885–906. https://doi.org/10.1037/a0031918

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science, 349*(6251). https://doi.org/10.1126/science.aac4716

Osinsky, R., Karl, C., & Hewig, J. (2017). Dispositional anxiety and frontal-midline theta: On the modulatory influence of sex and situational threat. *Journal of Personality, 85*(3), 300–312. https://doi.org/10.1111/jopy.12241

Peirce, J. W. (2007). PsychoPy—Psychophysics software in Python. *Journal of Neuroscience Methods, 162*(1-2), 8–13. https://doi.org/10.1016/j.jneumeth.2006.11.017

Pugnaghi, G., Cooper, A., Ettinger, U., & Corr, P. J. (2018). The psychometric properties of the German language reinforcement sensitivity theory-personality questionnaire (RST-PQ). *Journal of Individual Differences, 39*(3), 182–190.

Rammstedt, B., Kemper, C. J., Klein, M. C., Beierlein, C., & Kovaleva, A. (2012). *Eine kurze Skala zur Messung der fünf Dimensionen der Persönlichkeit: Big-Five-Inventory-10 (BFI-10)*. Mannheim, Germany: GESIS – Leibniz-Institut für Sozialwissenschaften.

Rodrigues, J., Ulrich, N., & Hewig, J. (2015). A neural signature of fairness in altruism: A game of theta? *Social Neuroscience, 10*(2), 192–205. https://doi.org/10.1080/17470919.2014.977401

Rodrigues, J., Ulrich, N., Mussel, P., Carlo, G., & Hewig, J. (2017). Measuring prosocial tendencies in Germany: Sources of validity and reliability of the revised prosocial tendency measure. *Frontiers in Psychology, 8*, Article 2119. https://doi.org/10.3389/fpsyg.2017.02119

Rodrigues, J., Weiß, M., Hewig, J., & Allen, J. J. B. (2021). EPOS: EEG processing open-source scripts. *Frontiers in Neuroscience, 16*, Article 660449. https://doi.org/10.3389/fnins.2021.660449

Sato, A., Yasuda, A., Ohira, H., Miyawaki, K., Nishikawa, M., Kumano, H., & Kuboki, T. (2005). Effects of value and reward magnitude on feedback negativity and P300. *Neuroreport, 16*(4), 407–411. https://doi.org/10.1097/00001756-200503150-00020

Schönbrodt, F. D., & Perugini, M. (2013). At what sample size do correlations stabilize? *Journal of Research in Personality, 47*(5), 609–612. https://doi.org/10.1016/j.jrp.2013.05.009

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science, 22*(11), 1359–1366. https://doi.org/10.1177/0956797611417632

Simonsohn, U., Simmons, J. P., & Nelson, L. D. (2020). Specification curve analysis. *Nature Human Behaviour, 4*(11), 1208–1214.

Sproesser, G., Strohbach, S., Schupp, H., & Renner, B. (2011). Candy or apple? How self-control resources and motives impact dietary healthiness in women. *Appetite, 56*, 784–787.

Stahl, J., Gibbons, H., & Miller, J. (2010). Modeling single-trial LRP waveforms using gamma functions. *Psychophysiology, 47*(1), 43–56. https://doi.org/10.1111/j.1469-8986.2009.00878.x

Steegen, S., Tuerlinckx, F., Gelman, A., & Vanpaemel, W. (2016). Increasing transparency through a multiverse analysis. *Perspectives on Psychological Science, 11*(5), 702–712. https://doi.org/10.1177/1745691616658637

Strobel, A., Beauducel, A., Debener, S., & Brocke, B. (2001). Eine deutschsprachige Version des BIS/BAS-Fragebogens von Carver und White [A German version of Carver and White's BIS/BAS scales]. *Zeitschrift für Differentielle und Diagnostische Psychologie, 22*(3), 216–227. https://doi.org/10.1024/0170-1789.22.3.216

Szucs, D., & Ioannidis, J. P. (2017). Empirical assessment of published effect sizes and power in the recent cognitive neuroscience and psychology literature. *PLoS Biology, 15*(3), Article e2000797. https://doi.org/10.1371/journal.pbio.2000797

Verbruggen, F., & Logan, G. D. (2008). Automatic and controlled response inhibition: Associative learning in the go/no-go and stop-signal paradigms. *Journal of Experimental Psychology. General, 137*(4), 649–672. https://doi.org/10.1037/a0013170

Wacker, J. (2017). Increasing the reproducibility of science through close cooperation and forking path analysis. *Frontiers in Psychology, 8*, Article 1332. https://doi.org/10.3389/fpsyg.2017.01332

Werner, R., & von Collani, G. (2004). Deutscher Aggressionsfragebogen. *Zusammenstellung sozialwissenschaftlicher Items und Skalen (ZIS)*. https://doi.org/10.6102/zis52

Wiltink, J., Vogelsang, U., & Beutel, M. E. (2006). Temperament and personality: The German version of the Adult Temperament Questionnaire (ATQ). *GMS Psycho-Social Medicine, 3*, Article Doc10.

World Health Organization (1998) Well-being measures in primary health care: The DepCare Project. Health for All. Target 12. Geneva, Switzerland: WHO.

Zimmermann, J., Altenstein, D., Krieger, T., Holtforth, M. G., Pretsch, J., Alexopoulos, J., Spitzer, C., Benecke, C., Krueger, R. F., Markon, K. E., & Leising, D. (2014). The structure and correlates of self-reported DSM-5 maladaptive personality traits: Findings from two German-speaking samples. *Journal of Personality Disorders, 28*(4), 518–540. https://doi.org/10.1521/pedi_2014_28_130

eapp

*Personality Science* (PS) is an official journal of the European Association of Personality Psychology (EAPP).

leibniz-psychology.org

PsychOpen GOLD is a publishing service by Leibniz Institute for Psychology (ZPID), Germany.